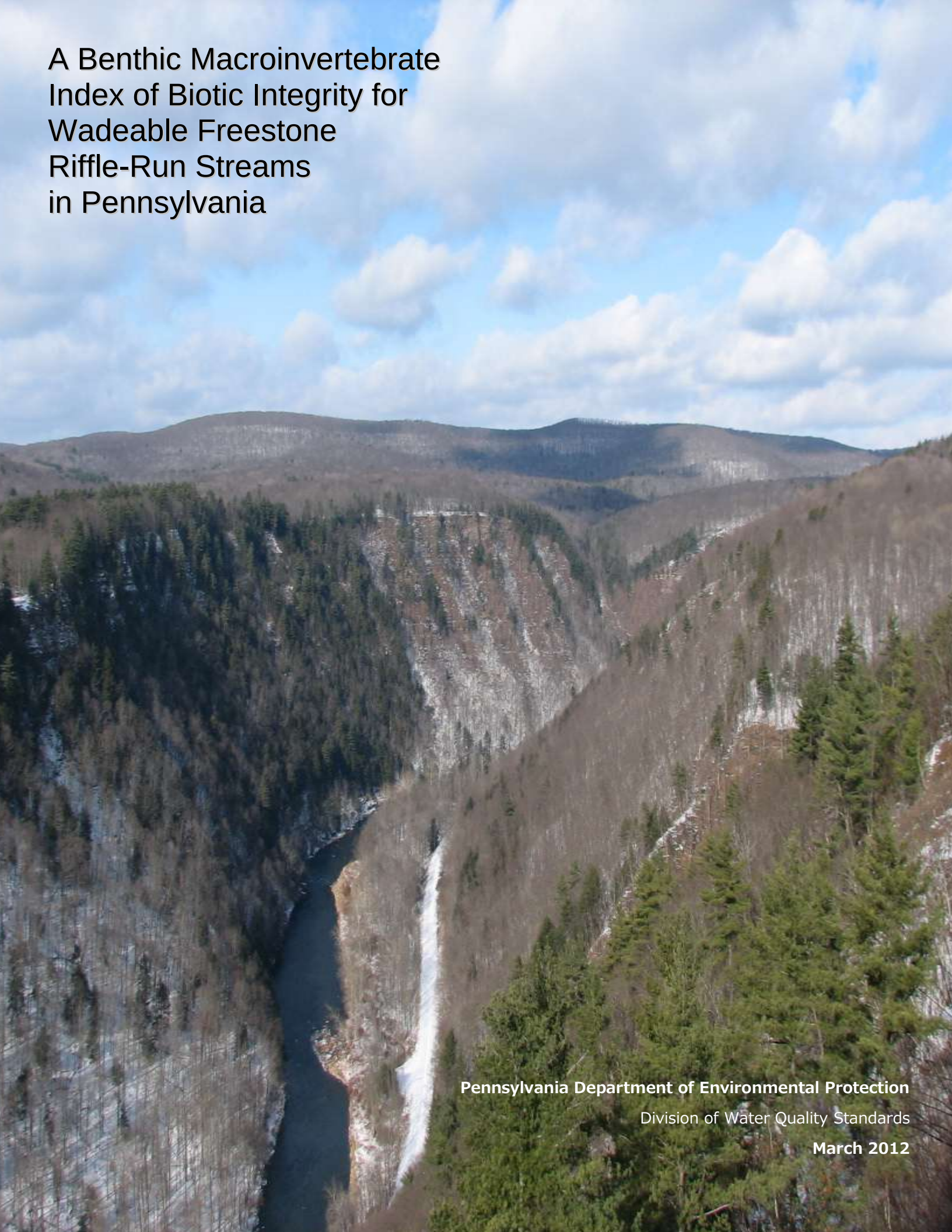
A Benthic Macroinvertebrate Index of Biotic Integrity for Wadeable Freestone Riffle-Run Streams in Pennsylvania



Pennsylvania Department of Environmental Protection

Division of Water Quality Standards

March 2012

PLEASE PRINT RESPONSIBLY!

Some of the figures and tables in this report require full color for comprehension and are best viewed in electronic format. In an effort to help conserve resources, please print only as much of this report as you really, really need to.

THANK YOU!

ACKNOWLEDGEMENTS

This project would not have been possible without the skills and dedication of biologists currently and formerly employed with the Pennsylvania Department of Environmental Protection and affiliated organizations. These wonderful people collected and processed the many thousands of fascinating organisms that form the foundation of this project:

Bill Andrus	**Tim Daley**	**Ron Hughey**	**John Ryder**
Kristen Bardell	**Jared Dressler**	**Gary Kenderes**	**Rob Ryder**
Steve Barondeau	**Scott Dudzic**	**Rod Kime**	**Tony Shaw**
Heidi Biggs	**Mark Embeck**	**Andy Klinger**	**Derek Smith**
Dan Bogar	**Alan Everett**	**Sherry Leap**	**Rick Spear**
Bill Botts	**Ed Filip**	**Kim Long**	**Kay Spyker**
Mike Boyer	**Aaron Frey**	**Josh Lookenbill**	**Olyssa Starry**
Mark Brickner	**Martin Friday**	**Rod McAllister**	**Harry Vitolins**
Joe Brancato	**Jay Gerber**	**Charlie McGarrell**	**Gary Walters**
Angela Bransteitter	**Joy Gillespie**	**Steve Means**	**Rick Weber**
Mark Brickner	**Jim Grazio**	**Eric Mosbacher**	**Carrie Wengert**
Brian Chalfant	**Joe Hepp**	**Abbey Owoc**	**Allen Whitehead**
Dan Counahan	**Jennifer Hill**	**Molly Pulket**	**Amy Williams**

Mike Bilger with EcoAnalysts, Inc.; **Tom Shervinskie** with the Pennsylvania Fish and Boat Commission; **Erik Silldorf** with the Delaware River Basin Commission; **Adam Griggs** with the Interstate Commission on the Potomac River Basin; **Jen Hoffman** and **Susan Buda** with the Susquehanna River Basin Commission; **Theodore Buckwalter**, **Heather Eggleston**, **Leif Olson**, **Andrew Reif** and **Mark Roland** with the United States Geological Survey; **Amy Seidel** with the Monroe County Planning Commission; **Celina Seftas** with Huntingdon County Conservation District; and **Doug Ebert** with Erie County Department of Health also collected and/or processed samples used in this project.

Andy Klinger and **Charlie McGarrell**, both formerly with the Pennsylvania Department of Environmental Protection's Central Office, provided helpful data analyses and advice, respectively, during the early phases this project. The members of the peer review committee for this project also provided valuable guidance and comments during its development. The committee members are listed in the table immediately below.

Name	Affiliation		Location
Bill Botts	Pennsylvania Department of Environmental Protection	Central Office	Harrisburg, PA
Rod Kime			
Tony Shaw			
Maggie Passmore	United States Environmental Protection Agency	Region III Office	Wheeling, WV
Greg Pond			
Karen Blocksom		Office of Research and Development	Cincinnati, OH
Mike Bilger	EcoAnalysts, Inc.		Selinsgrove, PA
John Jackson	Stroud Water Research Center		Avondale, PA
John Arway	Pennsylvania Fish and Boat Commission		Bellefonte, PA
Jeremy Deeds	Western Pennsylvania Conservancy		Pittsburgh, PA
Mary Walsh			Middletown, PA

The participants in the Pennsylvania Tiered Aquatic Life Use workshops, who also provided invaluable input and guidance for this project, are listed in the table immediately below.

Name	Workshop Year			Affiliation		Location
	2006	2007	2008			
Joe Brancato	X		X	Pennsylvania Department of Environmental Protection	Northwest Regional Office	Meadville, PA
Scott Dudzic	X	X				
Ron Hughey	X				Northcentral Regional Office	Williamsport, PA
Martin Friday	X	X	X			
Steve Means			X			
John Ryder			X			
Tim Daley	X	X			Northeast Regional Office	Wilkes-Barre, PA
Sherry Leap	X	X	X			
Abbey Falcone	X				Southwest Regional Office	Pittsburgh, PA
Rick Spear	X	X	X			
Kristen Bardell			X		Southcentral Regional Office	Harrisburg, PA
Mark Embeck	X					
Jay Gerber			X			
Joe Hepp		X	X			
Bob Schott	X	X	X			
Mike Boyer	X				Southeast Regional Office	Norristown, PA
Alan Everett	X	X	X			
Ed Filip			X			
Joy Gillespie			X			
Kim Long		X				
Dan Bogar	X	X	X		Central Office	Harrisburg, PA
Bill Botts	X	X	X			
Mark Brickner	X	X	X			
Angela Bransteitter		X	X			
Brian Chalfant	X	X	X			
Rod Kime	X	X	X			
Josh Lookenbill			X			
Charlie McGarrell	X					
Steve Moyer			X			
Jen Orr		X	X			
Molly Pulket		X	X			
Rob Ryder	X	X	X			
Tony Shaw		X				
Olyssa Starry	X					
Gary Walters	X		X			
Amy Williams			X			
Mary Walsh	X			Western Pennsylvania Conservancy		Middletown, PA
Jeroen Gerritsen	X			Tetra Tech, Inc.		Fairfax, VA
Benjamin Jessup	X					
Evan Horning	X			United States Environmental Protection Agency	Office of Water	Washington, DC
Maggie Passmore	X	X	X		Region III Office	Wheeling, WV
Greg Pond	X	X	X			
Mike Bilger	X	X	X	EcoAnalysts, Inc.		Selinsgrove, PA
Bob Limbeck			X	Delaware River Basin Commission		West Trenton, NJ
Erik Silldorf			X			

Thanks everybody!

TABLE OF CONTENTS

PROJECT SUMMARY

The principal motivation for this project was to develop an index of biological integrity (IBI) for benthic macroinvertebrate communities in Pennsylvania's *larger* wadeable, freestone, riffle-run streams. This project builds on previous work to develop a benthic macroinvertebrate IBI for *smaller* wadeable, freestone, riffle-run streams in Pennsylvania. The following report synthesizes analyses of benthic macroinvertebrate samples from wadeable, freestone, riffle-run streams across Pennsylvania – sizeable, tiny, and otherwise.

The IBI developed in this project incorporates six biological metrics that measure relevant aspects of benthic macroinvertebrate community composition in Pennsylvania's wadeable, freestone, riffle-run streams. Before combining the individual metrics into IBI scores, two different sets of metric standardization values are applied. One set of metric standardization values is applied to samples from smaller streams while a second set of values is applied to samples from larger streams. Broadly speaking, smaller streams are characterized as first through third order streams (using the Strahler stream ordering system) that drain 25 or fewer square miles of land. For this project, larger streams are broadly characterized as fifth and higher order *wadeable* streams draining 50 or more square miles of land – different sampling and assessment protocols apply to *non-wadeable* rivers. Detailed discussion about how to apply these procedures as well as considerations about whether to apply the large-stream or the small-stream procedures to fourth order freestone streams and streams draining 25 to 50 square miles are discussed in detail in the body of this report.

Aquatic life use attainment benchmarks are established based on IBI scores. Different benchmarks apply to samples collected in different seasons. One set of benchmarks applies to samples collected from November to May and another set of benchmarks applies to samples collected from June to September. Depending on the particular climatological conditions in a given year and other considerations discussed in this report, either of these two sets of seasonal benchmarks can apply to samples collected during October. Different benchmarks and assessment criteria are also developed for streams with different protected aquatic life uses. To strengthen the assessment process, a series of additional biological screening criteria – detailed in the report – are applied to samples from streams of different sizes at different times of the year.

The biological and ecological concepts concerning changes in the composition of benthic macroinvertebrate communities related to stream size (e.g., Vannote et al. 1980) and annual seasons are well established. This project provides analyses that support specific stream size and seasonal classifications for an IBI and aquatic life use assessment procedures for benthic macroinvertebrate samples from Pennsylvania's wadeable, freestone, riffle-run streams. This report also presents some considerations for applying the index to wadeable, *limestone-influenced*, riffle-run streams. Separate protocols exist for evaluating lower gradient pool-glide streams (PADEP 2007) as well as true limestone spring streams (PADEP 2009a).

Happy reading!

INTRODUCTION

This project aims to develop an indicator of biological integrity for benthic macroinvertebrate communities in the wadeable, freestone, riffle-run streams of Pennsylvania. Through direct quantification of biological attributes along a gradient of ecosystem conditions, this indicator will measure the extent to which anthropogenic activities compromise a stream's ability to support healthy aquatic communities (Davis and Simon 1995). This biological assessment tool will help guide and evaluate legislation, policy, goals, and management strategies for Pennsylvania's aquatic resources (Davis and Simon 1995; Davies and Jackson 2006; Hawkins 2006).

Legislative Background

The objective of the United States Federal Water Pollution Control Act (United States Code 2011: Title 33, Sections 1251 through 1387) – more commonly known as the Clean Water Act – as stated in section 1251(a) is,

> "to restore and maintain the chemical, physical, and biological integrity of the Nation's waters."

An interim goal of the Clean Water Act as stated in Section 1251(a)(2) is,

> "… water quality which provides for the protection and propagation of fish, shellfish, and wildlife…"

Section 1251(b) of the Clean Water Act indicates that the primary authority and responsibility for prevention, reduction, and elimination of pollution as well as for management of land and water resources rests with the States. Thus, States are responsible for setting water quality goals to protect aquatic life. To this end, States have defined various levels of designated aquatic life use (ALU) – such as recreational fishing and fish migration – to be protected for specific water bodies.

In addition to the federal Clean Water Act, Pennsylvania's Clean Streams Law (35 P. S. § § 691.1 – 691.1001) aims to,

> "… preserve and improve the purity of the waters of the Commonwealth for the protection of public health, animal and aquatic life, and for industrial consumption, and recreation…"

To this end, the Pennsylvania Code (2011: Title 25, Chapter 93.3) recognizes four categories of protected ALUs, including: (1) cold water fishes (CWF); (2) warm water fishes (WWF); (3) migratory fishes (MF); and (4) trout stocking (TSF). The CWF and WWF uses include protection of fish as well as additional flora and fauna (e.g., benthic macroinvertebrates) indigenous to a cold or warm water habitat, respectively. The TSF use also includes protection of fish and additional flora/fauna indigenous to a warm water habitat. Pennsylvania regulations also recognize two antidegradation – or "special protection" – water uses: high quality waters (HQ) and exceptional value waters (EV). Details concerning these uses and their application to Pennsylvania's waters can be found

in Chapter 93 of the Pennsylvania Code.

Biological Monitoring

To meet the objectives outlined in the federal Clean Water Act – as well as Pennsylvania's Clean Streams Law – evaluations of aquatic ecosystem integrity ideally include evaluations of physical characteristics (e.g. types and distribution of habitats and substrates; flow patterns; channel stability), water chemistry (e.g., concentrations of toxic and nontoxic chemicals), and biological communities (e.g., fish, benthic macroinvertebrates, periphyton). However, chemical water quality evaluations are of limited value in assessing overall ecosystem condition because of the difficulty of evaluating every relevant chemical parameter, the synergistic chemical effects on ecosystems, and the highly transient nature of lotic water chemistry, as well as cost and logistical considerations of frequent chemical monitoring. Abiotic physical evaluations of streams – although informative in many respects – are also of limited value in assessing overall ecosystem integrity for a wide array of stressors. For example, in some acid deposition situations, watershed and in-stream physical conditions may be largely undisturbed, but the biotic community may be drastically altered by the acidification.

Biological monitoring offers the ability to assess long-term, cumulative effects of many types of ecosystem stress, including stress related to chemical and physical habitat factors. Organisms living in aquatic environments are intimately associated with and affected by chemical water quality and the physical conditions of streams and watersheds. As such, these organisms can be viewed as living indicators of overall ecosystem condition. However, biological monitoring also has its limitations and cannot always unequivocally identify causative stressors, which may be better identified when biological data is viewed in conjunction with chemical water quality and physical habitat assessments (Novotny 2004).

Indicators of biological integrity – based on direct measures of community and population response – provide relevant and useful tools that can be used independently, or in concert with other information (e.g., physical and chemical evaluations) for the purpose of assessing protected ALUs (Novotny 2004).

Indicators of Biological Integrity

Although the Clean Water Act outlines the general objective of biological integrity, no legislation explicitly defines biological integrity. The United States House and Senate Committee on Public Works deliberations on the Clean Water Act included the concept of "naturalness" as a key part of biological integrity (see Stoddard et al. 2006). Legislation in the United States, Europe, and Australia expresses a need to characterize biological conditions that occur in natural states, with minimal human impacts (Stoddard et al. 2006).

Consistent with this concept, a definition of biological integrity proposed and endorsed by many ecologists states that an ecosystem with biological integrity supports and maintains a balanced, integrated, adaptive system having a full range of ecosystem elements (e.g., genes, species, assemblages) and processes (e.g., mutation, metapopulation dynamics, nutrient and energy dynamics) expected in areas with minimal human influence (Karr and Dudley 1981; Davis and Simon 1995; Davies and Jackson 2006).

Monitoring and assessment of the biological integrity of inland water resources across the world frequently involves measuring the degree to which community-level biological

attributes (e.g., structure, composition, function, diversity) differ from a community minimally influenced by human activities: a reference community (Davis and Simon 1995; Davies and Jackson 2006; Hawkins 2006; Stoddard et al. 2006). Often, a major goal of biological monitoring and assessment is to describe the impacts of human activities on the structure and function of aquatic ecosystems (Stoddard et al. 2006).

Accurate assessment of biological condition requires integration of biological responses at varying scales, from individual organism responses to community-level responses and ecosystem-level responses (Barbour et al. 1995). Past efforts have helped develop and refine the science of using biological indicators to assess ecosystem conditions (Hawkins 2006). Such indicators of biological integrity help to document environmental conditions at community and ecosystem levels, which can assist in diagnostic analyses of sources and causes of ecosystem stress (Barbour et al. 1995).

Many States have developed and are using indicators of biological integrity based on stream benthic macroinvertebrate communities as ALU assessment tools, including Maryland (Stribling et al. 1998), West Virginia (Gerritsen et al. 2000), Virginia (Burton and Gerritsen 2003), and Kentucky (Pond et al. 2003) among many others.

The Commonwealth and Its Waters

The Commonwealth of Pennsylvania encompasses approximately 45,000 square miles of land (Figure 1) with diverse climatic, geological, physiographic, and land use characteristics. Well over 80,000 miles of flowing waters drain Pennsylvania's varied landscape, ranging from ephemeral headwater hollows, small perennial creeks and brooks, to massive rivers such as the Ohio, Delaware, and Susquehanna.

The Pennsylvania Department of Environmental Protection (PADEP) recognizes that certain types of streams naturally differ in physiochemical, climatological, geological, and many other -ological characteristics and, consequently, in biological potential. For example, benthic macroinvertebrate communities in limestone spring streams (streams in which most or all of the flow arises from springs and groundwater in areas with primarily calcareous geologies) often exhibit noticeably different characteristics (e.g. low diversity, high abundance) than communities in many freestone streams. These differences are attributable, in large part, to the unique physiochemical conditions associated with spring-fed, groundwater-dominated streams (e.g., relatively constant thermal and flow regimes).

Currently, PADEP utilizes three different methodologies to monitor and assess the benthic macroinvertebrate communities in three types of streams in Pennsylvania: true limestone spring streams (PADEP 2009a); lower gradient pool-glide type streams (PADEP 2007); and wadeable, freestone, riffle-run type streams. The last of these three stream types is the focus of this project. PADEP is also currently developing biological assessment methods for large, non-wadeable rivers.

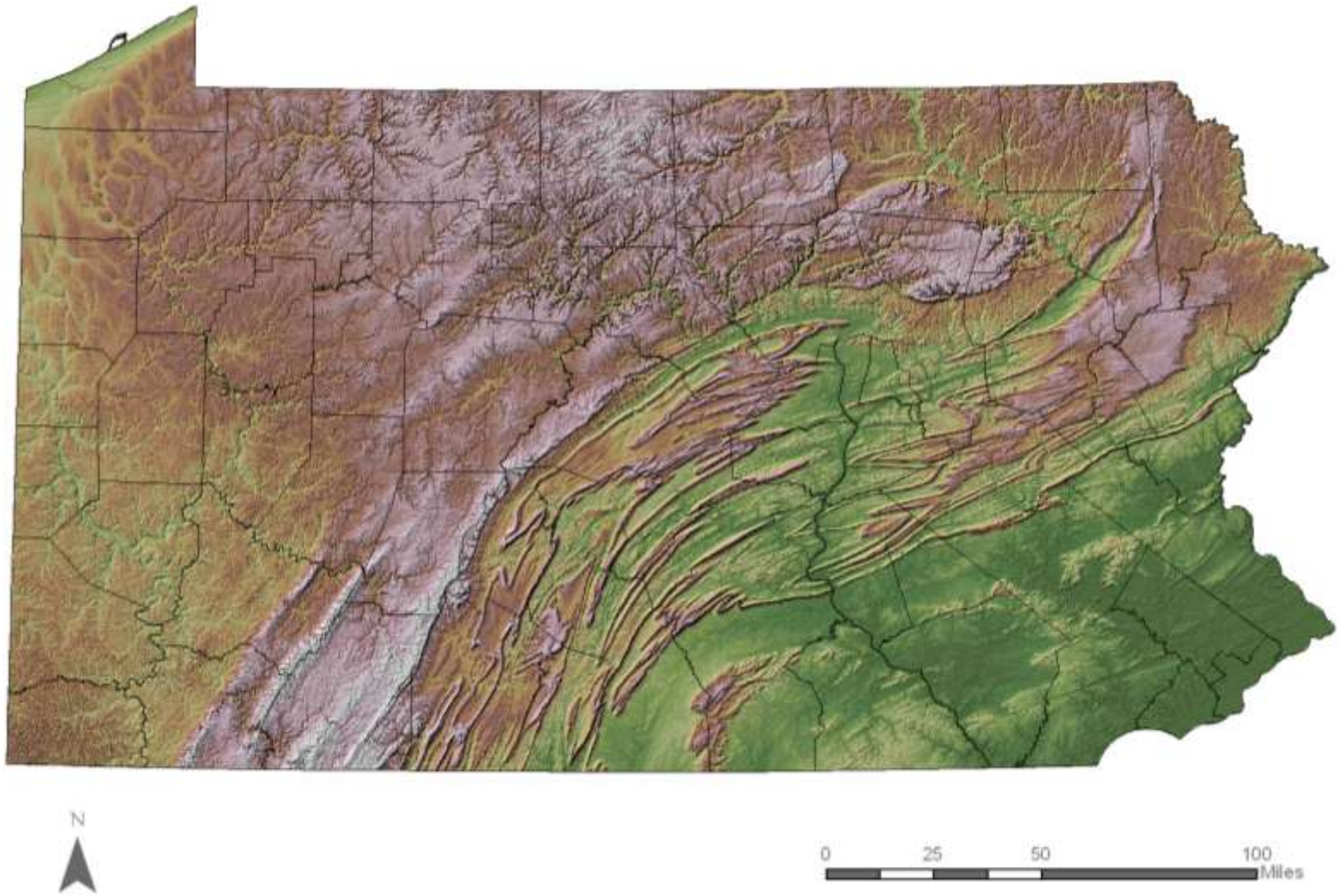


Figure 1. Shaded relief map of the Commonwealth of Pennsylvania, with county boundaries.

DATA COLLECTION METHODS

All benthic macroinvertebrate samples analyzed in this project were collected using D-frame nets with 500-micron mesh. At a sampling site, biologists worked progressively upstream, compositing six kicks from riffle areas distributed throughout a 100-meter stream reach. Biologists sampled areas representative of the variety of riffle habitats (e.g., slower flowing, shallow riffles and faster flowing, deeper riffles) present within the sample reach. With each kick, biologists disturbed approximately one square meter immediately upstream of the net for approximately one minute to an approximate depth of 10 cm, as substrate allowed. Composited samples were preserved with 95% ethanol in the field and transported back to the laboratory for processing.

In the lab, each composited sample was placed into a 3.5" deep rectangular pan (measuring 14" long x 8" wide on the bottom of the pan) marked off into 28 four-square inch (2" x 2") grids. Four of the grids were randomly selected. The contents of the randomly selected grids were extracted – using plastic spoons, knives, turkey basters, and other implements as needed – from within four-square inch circular "cookie cutters" placed in the selected grids in the pan. These extracted contents were then placed into a second pan with the same dimensions and markings as the initial pan. All the organisms were picked from this second pan.

If less than 160 identifiable organisms were picked from the second pan, an additional grid was randomly selected and extracted from the first pan. The contents of this additional grid were transferred to the second pan, and the organisms were picked from the second pan. This process was continued until the target number of organisms was reached. The target number of organisms was 200 ± 40 identifiable organisms, with 190 to 210 identifiable organisms being the preferred range. In situations with a count of identifiable organisms in a sub-sample between 160 and 180 and a sample that has not been entirely picked, PADEP highly encourages picking an additional grid or two to get closer to the target number of 200 identifiable organisms (i.e., in the preferred 190 to 210 organism range).

If more than 240 identifiable organisms were picked from the initial four grids, then those organisms were all placed into another pan and floated. A grid was then randomly selected and the organisms were picked from the selected grid. This process continued until the target number of organisms (200 ± 40, with 190 to 210 preferred) was reached.

Any grid selected during any part of the sub-sampling process was picked in its entirety. The total number of grids selected for each part of the sub-sampling process (e.g., 4 of 28 grids from the first pan, 10 of 28 grids from the second pan) was recorded.

Organisms in the sub-sample were identified and counted. Midges were identified to the family level of Chironomidae. Snails, clams, and mussels were all also identified to family levels. Roundworms and proboscis worms were identified to the phylum levels of Nematoda and Nemertea, respectively. Moss animacules were identified to the phylum level of Bryozoa. Flatworms and leeches were identified to the class levels of Turbellaria and Hirudenia, respectively. Segmented worms, aquatic earthworms, and tubificids were identified to the class level of Oligochaeta. All water mites were identified as Hydracarina, an artificial taxonomic grouping of several mite superfamilies. All other macroinvertebrates

were identified to genus level. Field sampling and laboratory methods are more fully described in Appendix A.

Land uses were calculated for the upstream basins of each sampling location using ESRI ® ArcMap™ 9.3 geographic information system (GIS) software and the 2001 National Land Cover Dataset (Homer et al. 2004).

Biologists collected water chemistry samples and conducted physical habitat assessments concurrently with many macroinvertebrate samples, although not all macroinvertebrate samples in the dataset had accompanying water chemistry and habitat data.

In addition to benthic macroinvertebrates, land use, water chemistry, and physical habitat data, a suite of GIS-based data were included in the analysis for each sample, including: watershed area; Strahler stream order; river basin; county; sampling location elevation; current ALU and attainment status of the stream segment from which the sample was taken; proportion of stream miles upstream impaired by various sources and causes; geologic composition of the watershed; and slope of the stream segment from which the sample was taken. Strahler stream order was determined from the 1:100,000-scale National Hydrography Dataset Plus (http://www.horizon-systems.com/nhdplus) and from an internal PADEP GIS stream layer. Slope data was derived from Anderson and Olivero (2003) and from Gawler et al. (2008).

Numerous biologists (see Acknowledgements) collected the data used in this analysis. The samples in the dataset were collected for a variety of PADEP survey types, with most samples collected as part of in-stream comprehensive evaluation surveys (1,167 samples) and antidegradation surveys (773 samples). Some samples in this dataset were also collected as probabilistic surveys (341 samples), long-term fixed-site water quality network monitoring surveys (264 samples), cause and effect surveys (186 samples), effluent dominated stream surveys (127 samples), intensive unassessed follow-up surveys (48 samples), basin surveys (46 samples), benthic macroinvertebrate surveys at fish sampling sites (38 samples), use attainability surveys (34 samples), point of first use surveys (14 samples), nonpoint source remediation surveys (5 samples), outside agency surveys (4 samples), and a stream enrichment risk analysis survey (1 sample).

In areas with multiple samples taken within a short distance (i.e., within a few hundred meters on the same stream reach), nearby samples were considered to be from one site, unless there were major intervening differences between spatially proximate samples (e.g., samples collected just upstream and just downstream of a discharge; substantial changes in land use between samples), in which case nearby samples were considered as representing distinct sites.

SITES AND SAMPLES

The dataset consisted of 3,047 benthic macroinvertebrate samples from 2,480 sites. All sites were located within the borders of the Commonwealth of Pennsylvania except for five sites on larger streams in the Potomac River basin whose headwaters are in Pennsylvania (Figure 2). Samples from these five Potomac basin sites were collected at long-term, fixed-location monitoring sites located just south of the Mason-Dixon Line in Maryland on Antietam Creek, Conococheague Creek, Tonoloway Creek, Town Creek, and Sideling Hill Creek.

Although samples were collected from sites representing many areas of Pennsylvania, some basins had noticeably higher sampling densities than other basins (Figure 2) as a result of PADEP's rotating basin monitoring strategy.

In terms of major basins in Pennsylvania, sampling densities were particularly high in the following basins:

- Brandywine River - Christina River
- Lower West Branch Susquehanna River
- Lehigh River
- Middle Delaware River
- Middle Allegheny River - Tionesta Creek
- Chautauqua Creek - Conneaut Creek
- Clarion River
- Schuylkill River
- Lackawaxen River

Sampling densities were also high in some basins that drain smaller areas of Pennsylvania such as:

- Upper Genesee River
- Gunpowder River - Patapsco River
- Monocacy River
- Upper Monongahela River
- Crosswicks Creek - Neshaminy Creek

Of the larger basins in Pennsylvania, sampling densities were lowest in the following basins:

- Upper Ohio River
- Raystown Branch Juniata River
- Upper West Branch Susquehanna River
- Conemaugh River
- Lower Juniata River
- Upper Susquehanna River - Lackawanna River
- Connoquenessing Creek
- Lower Delaware River
- Upper Allegheny River
- Lower Monongahela River
- Shenango River

Sampling densities were also low in some basins that drain smaller areas of Pennsylvania such as:

- Owego Creek - Wappasening Creek
- North Branch Potomac River
- Mahoning River
- Cacapon River - Town Creek
- Cheat River

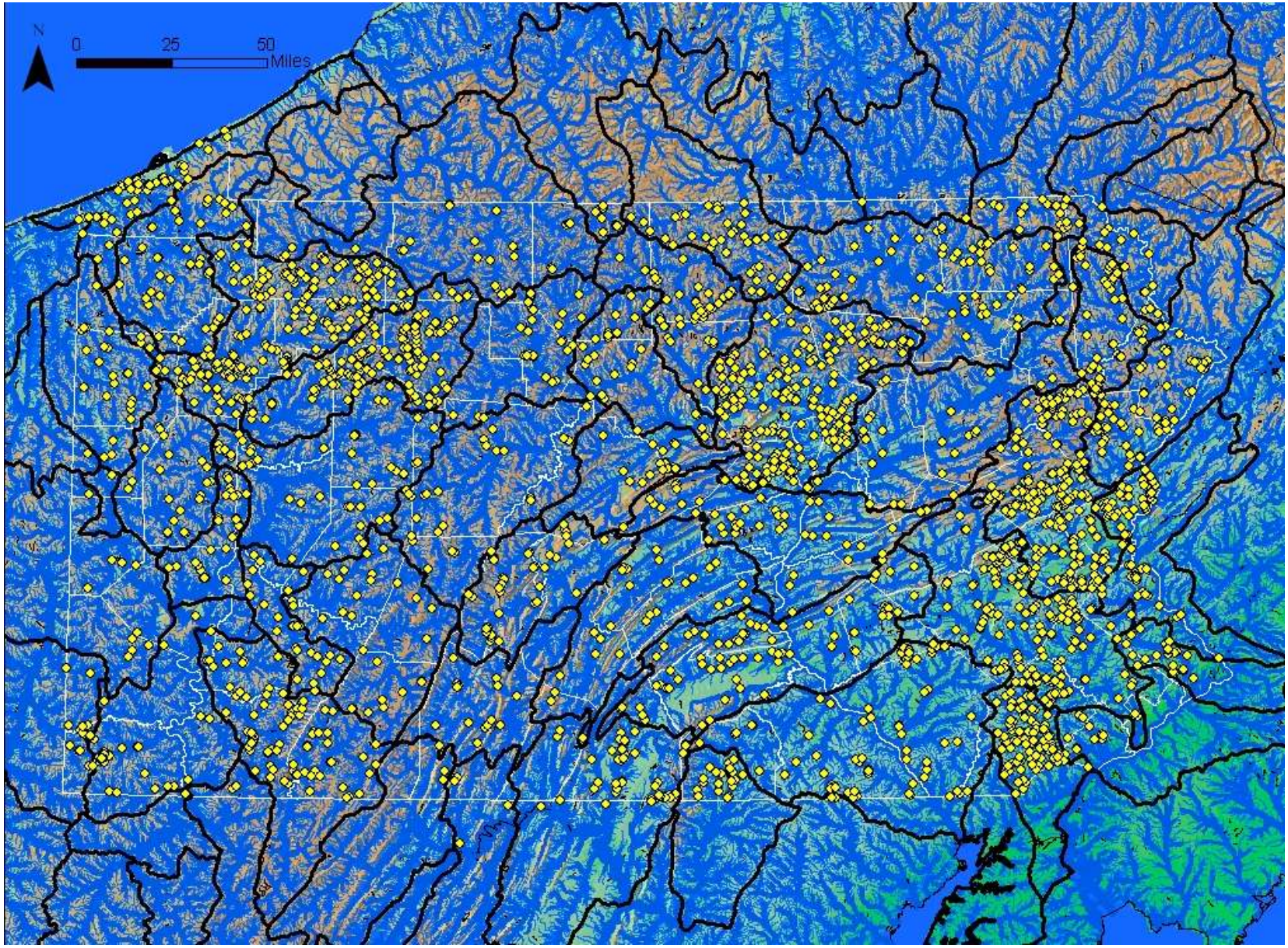


Figure 2. Sample site locations, with larger streams, Pennsylvania county boundaries, and major watershed boundaries.

Most of the samples were collected from first through third Strahler order stream reaches draining less than 25 square miles of land (Table 1).

Table 1. Number of samples by drainage area ranges and Strahler stream order.

Drainage area range (square miles)	Strahler stream order							
	1	2	3	4	5	6	7	8
0 to 3	364	551	70					
3 to 10	15	348	433	31				
10 to 25		12	323	172	10			
25 to 50		2	47	149	16			
50 to 100			2	106	63			
100 to 500			1	20	185	50		
500 to 1,000					5	44	1	
1,000 to 5,000					1	4	13	
5,000 to 10,000							4	2
> 10,000							1	2

Samples were collected from streams at a range of elevations (Figure 3, Figure 4) with a range of slopes (Figure 4, Figure 5). The smallest stream sites tended to have the highest slopes while larger stream sites tended to have lower slopes (Figure 5).

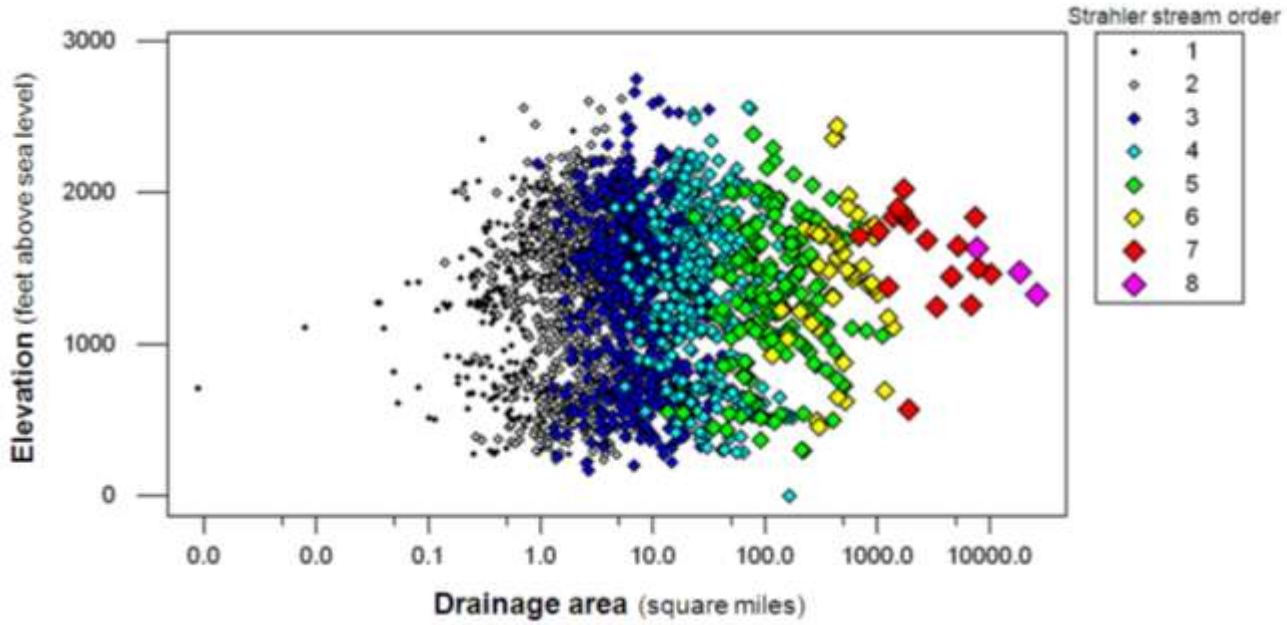


Figure 3. Relationship of sample site elevation and drainage area coded by Strahler stream order. Note logarithmic scale for drainage area.

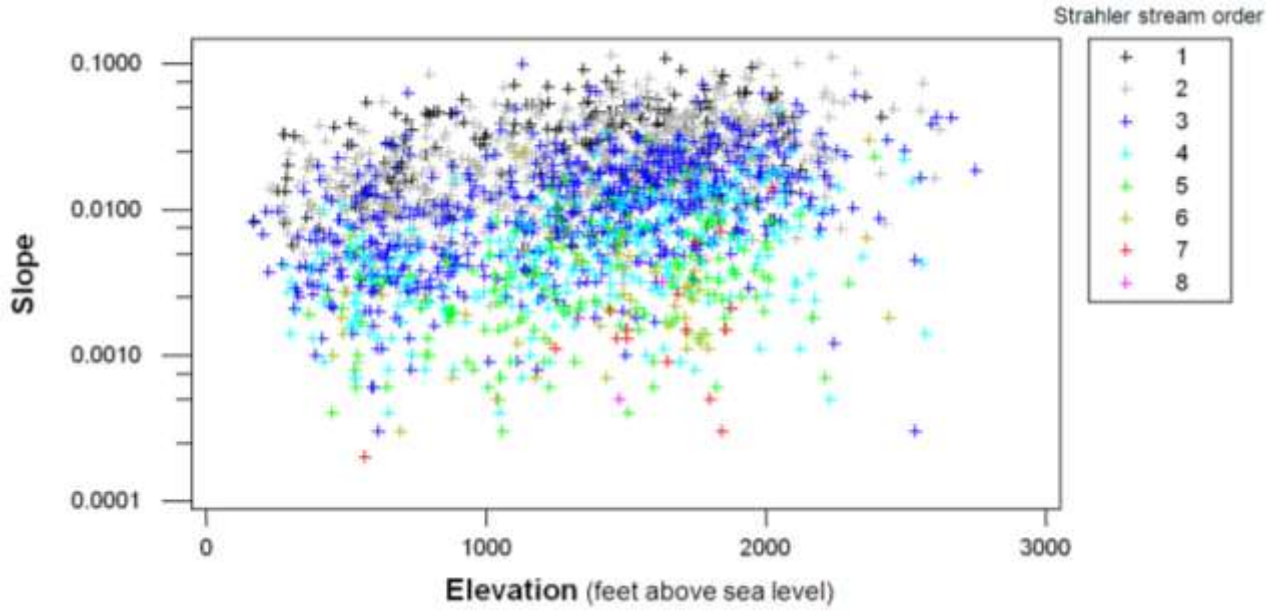


Figure 4. Relationship of sample site slope and elevation coded by Strahler stream order. Note logarithmic scale for slope. This figure only includes 2,690 samples for which slope data was readily available. Slope is presented as a ratio of vertical drop over longitudial distance.

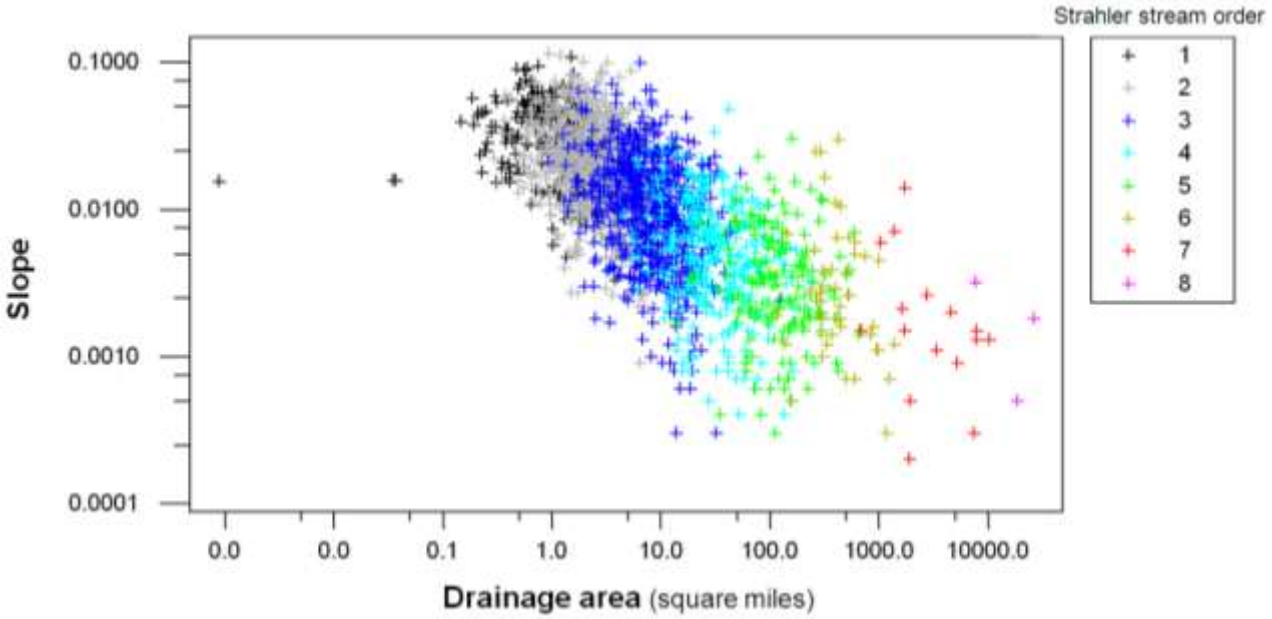


Figure 5. Relationship of sample site slope and drainage area coded by Strahler stream order. Note logarithmic scales on both axes. This figure only includes 2,692 samples for which slope and drainage area data were readily available. Slope is presented as a ratio of vertical drop over longitudial distance.

Samples were collected from November 10, 1999 to June 4, 2010 with about 75% of samples collected in 2006, 2007, 2008, and 2009 (Figure 6, Table 2). Around 60% of samples were collected during the months of March, April, and May (Figure 7, Table 2). Smaller stream sites tended to be sampled proportionally more in the spring while the largest stream sites tended to be sampled more in late summer and autumn (Figure 8).

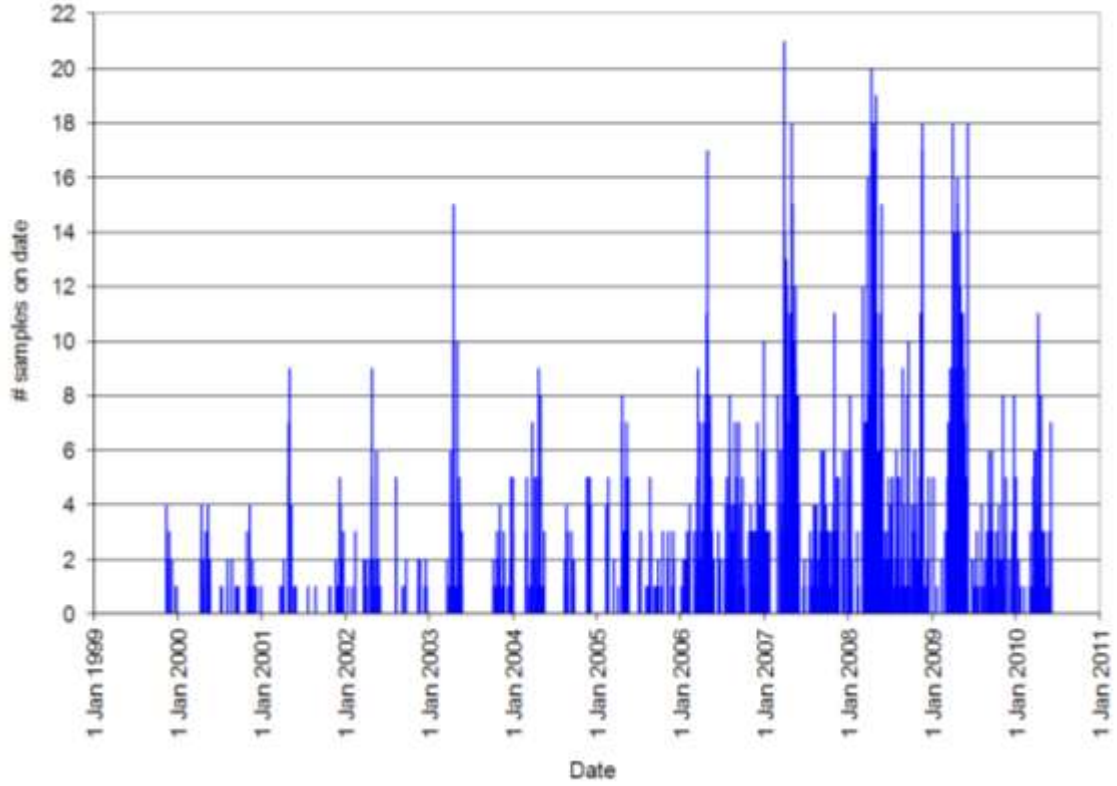


Figure 6. Distribution of samples by sampling date.

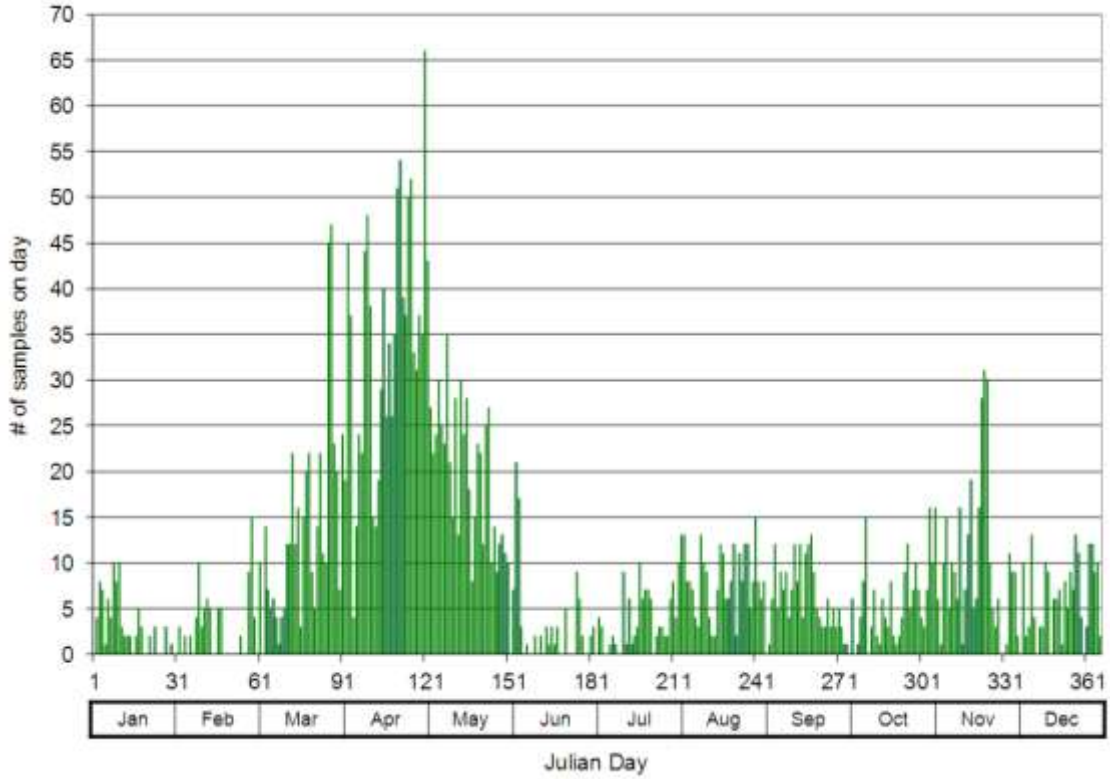


Figure 7. Distribution of samples by Julian day of sample collection.

Table 2. Sample collection dates by month and year.

Month	Year												# of samples by month	% of samples by month
	1999	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010		
January				2				10	22	38	5	5	**82**	**2.7%**
February				3		9	22	13	15	4	9	1	**76**	**2.5%**
March			2	5	3	14	2	50	92	85	100	31	**384**	**12.6%**
April		14	14	24	53	55	38	99	120	242	170	65	**894**	**29.3%**
May		15	22	24	32	4	25	44	148	124	101	11	**550**	**18.1%**
June							2	7	6	26	29	9	**79**	**2.6%**
July		2	1				3	30	12	23	6		**77**	**2.5%**
August		5	1	8		22	15	58	39	34	11		**193**	**6.3%**
September		4		5		10	6	22	39	29	36		**151**	**5.0%**
October		4	2		14		6	16	28	39	15		**124**	**4.1%**
November	12	10	4	5	11	19	12	41	36	94	21		**265**	**8.7%**
December	4	2	19	5	23	4	2	40	20	13	40		**172**	**5.6%**
# of samples by year	**16**	**56**	**65**	**81**	**136**	**137**	**133**	**430**	**577**	**751**	**543**	**122**	**3,047**	
% of samples by year	**0.5%**	**1.8%**	**2.1%**	**2.7%**	**4.5%**	**4.5%**	**4.4%**	**14.1%**	**18.9%**	**24.7%**	**17.8%**	**4.0%**		

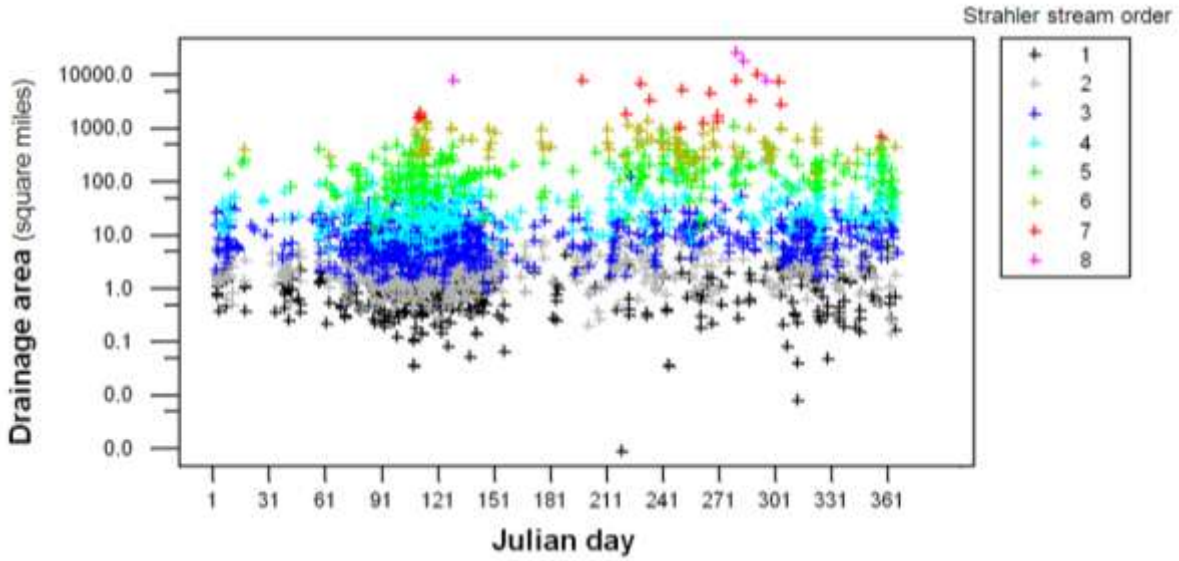


Figure 8. Distribution of samples by drainage area and Julian day, coded by Strahler stream order. Note logarithmic scale for drainage area.

Just under half of the samples in the dataset were collected from stream segments with protected antidegradation ALUs of EV or HQ, with about 25% of samples collected from streams with protected CWF, 9% with TSF, and 17% with WWF ALUs (Table 3).

Table 3. Number of samples by stream order and protected aquatic life use. Protected uses were undetermined for 17 samples (< 1% of the total number of samples) because the stream segments were not digitized in the National Hydrography Dataset.

Strahler stream order	Protected aquatic life use						
	EV	HQ-CWF	HQ-TSF	HQ-WWF	CWF	TSF	WWF
1	75	123	9	16	89	13	44
2	198	234	20	7	264	58	130
3	181	279	18	5	219	71	103
4	110	131	10	4	106	48	68
5	17	26	8	3	57	59	108
6	12	2	25		15	10	32
7					4		15
8							4
total #	**593**	**795**	**90**	**35**	**754**	**259**	**504**
% of total	**19%**	**26%**	**3%**	**1%**	**25%**	**9%**	**17%**

DEFINING SITE CONDITION

A critical step in development and implementation of any indicator of biological integrity used to evaluate effects of human activities on stream ecosystems involves quantification and comparison of the current condition of a stream's biology to a standard or benchmark condition. The standard or benchmark condition is often referred to as the reference condition and can be defined for a given type of water body and a given ALU (Hughes 1995; Barbour et al. 1999; Hawkins 2006; Stoddard et al. 2006). This reference condition represents the desired state of biotic assemblages based on relatively undisturbed conditions representative of a region and serves as the foundation for development of biological criteria (Hughes 1995; Stoddard et al. 2006). Reference conditions must be tailored to certain regions or certain types of water bodies because attainable biological conditions cannot be expected to be the same for every region or type of water body. For example, one would expect naturally different biological conditions in a stream in a tropical rainforest than in an arctic lake. The reference condition is usually defined as a range of conditions resulting from natural temporal and spatial variation and sampling error (Hughes 1995; Stoddard et al. 2006).

Expectations of biological condition can be estimated in a number of ways, including: the reference site approach (i.e., comparison to minimally or least disturbed sites); best professional judgment; interpretation of historical conditions; extrapolation of empirical models; and evaluation of ambient distributions (Hughes 1995; Stoddard et al. 2006). Each method of determining the reference condition has its own strengths and weaknesses and each method relies on ecosystem classification to some degree (Hughes 1995). The most useful means of defining reference conditions draw on all these approaches (Hughes 1995).

Although the process of defining the reference condition should be as objective as possible (e.g., use of defined abiotic criteria), considerable professional judgment is involved in site selection, data analysis and subsequent determination of acceptable versus unacceptable indicator scores (Hughes 1995). Professional sagacity can be difficult to quantify, but it plays an important role in any method of defining the reference condition (Hughes 1995) and can be strengthened when used in concert with other methods, such as abiotic criteria. Experienced biologists can develop empirical understanding of biological conditions in the absence of substantial human disturbance (Stoddard et al. 2006). The scientific credibility of professional judgment improves if it is tied to sound ecological theory, can be replicated by similarly experienced peers, and any decision rules or guidelines can be documented or quantified (Stoddard et al. 2006). The discussion later in this paper about PADEP's tiered aquatic life use workshops further explores the scientific credibility of applying professional judgment to macroinvertebrate communities in the wadeable, freestone, riffle-run streams of Pennsylvania, with encouraging results.

Stoddard et al. (2006) argue that the term *reference condition* should be used consistently to refer to a state of naturalness of the biotic structure and function, and that "naturalness implies the absence of significant human disturbance or alteration." Stoddard et al. (2006) also propose that this reference condition should be properly referred to as the *reference condition of biological integrity*. Stoddard et al. (2006) define four additional terms to describe the expected condition to which current conditions are compared, including: (1)

minimally disturbed condition; (2) historical condition; (3) least disturbed condition; and (4) best attainable condition.

In many areas, if not all over the planet, it is difficult to locate sampling sites representative of the natural state, or reference condition of biological integrity, and the goal of "pristine" waters (i.e., free from all human impacts) is an unrealistic goal due to widespread human impacts. As a result, reference conditions and water resource goals often practically describe minimally disturbed, least disturbed, or best attainable conditions (Hughes 1995; Novotny 2004; Stoddard et al. 2006). However, it is important to select reference sites representative of a region and ecosystem type that are disturbed as little as possible by human activities because the definition of the reference site has important consequences for development of biological indicators and subsequent establishment of ALU attainment thresholds (Hughes 1995; Barbour et al. 1999).

For natural resource management purposes, defining the reference condition helps establish the ecological potential of aquatic ecosystem types in a region while accounting for irreversible and reversible changes caused by humans (Novotny 2004). Reference sites representing least-disturbed conditions are moving targets of which human activities and natural processes are a part (Hughes 1995; Stoddard et al. 2006), but the range of conditions defined by what Stoddard et al. (2006) name the *minimally disturbed condition* should serve as a nearly invariant anchor by which we can assess ecosystem integrity.

Limited resources, time and data often hinder our ability to holistically assess exposure of stream ecosystems to the full range of stressors that impact them, so suites of criteria are often used to describe the characteristics of sites in a region that are least and most exposed to stressors, representing reference and stressed conditions respectively (Stoddard et al. 2006).

This project defines a reference condition based on a population of sites exhibiting biological integrity from across Pennsylvania to which sites of unknown biological integrity can be compared (Hughes 1995). This population-based approach to defining reference conditions provides comparability of samples for sites across the state from similar types of water bodies (i.e., wadeable, freestone, riffle-run streams) and promotes efficient use of limited public resources for monitoring and assessment of aquatic resources.

For this project, a suite of abiotic parameters comprised of watershed land uses, physical habitat evaluations, abandoned mine land prevalence, upstream ALU impairments, and water chemistry was used to determine relative anthropogenic impacts at each site and to define reference conditions. Initial site condition categories were assigned with a site condition index calculated from metrics of upstream land use, physical habitat evaluations, abandoned mine land prevalence, and upstream ALU impairments. The components of the initial site condition index were calculated for the upstream basin at each site as follows:

Land use component = (% forest + % wetlands) –
(% high-density development * 5) –
(% medium-density development * 3) –
(% low-density development *2) –
(% row crops) –
(% hay or pasture * 0.5)

Physical habitat component = minimum total habitat score at site / 240 * 100

Abandoned mine lands component = (% abandoned mine lands * -2)

Upstream impairments component = (% impaired stream miles * -1)

These four components were added together to calculate the initial site condition index for each site.

Initial site condition index = Land use component +
Physical habitat component +
Abandoned mine lands component +
Upstream impairments component

As shown above, various weightings were applied to the land use, abandoned mine lands, and upstream impairments components of the condition index. A number of different site characterization approaches were evaluated. The component weighting and condition index approach presented above is based on empirical observation and reasoning about how different types of impacts affect streams and benthic macroinvertebrate communities. For example, relatively small areas of high-density development can cause severe impacts to a stream by drastically altering flow patterns (e.g., increased overland runoff associated with impervious surfaces). An equal spatial extent of hay or pasture often has much less pronounced in-stream effects. In other words, if three percent of an otherwise forested stream's watershed is densely developed and imperviously paved, this will often have a much more severe impact on the basin's streams than if that three percent of land were utilized for hay or pasture. That is why high-density developed land use percentage was given a weighting of five while hay/pasture land use percentage was assigned a weighting of one-half. Similar reasoning was used to assign the weightings for each site condition component. Of course, the impact of any human activity on a stream depends on where the activity is located in the basin relative to the stream and a host of other situation-specific factors. However – for the purposes of this project – the site condition index as presented above represents a useful, tenable tool for comparing watershed condition across a large number of sites. There are certainly instances where the index does not holistically gauge the condition of certain streams and watersheds, but – by and large – it accomplishes its intent of quantifying the level of anthropogenic impacts to streams and their basins. This multifaceted quantification of anthropogenic impacts is conceptually similar to that of the Ecological Risk Index developed by Mattson and Angermeier (2007).

The initial site condition index values ranged from a maximum of 197 to a minimum of -255 (Figure 9). Higher initial site condition index values represent relatively pristine watersheds and streams while lower values represent streams and watersheds more impacted by anthropogenic activities.

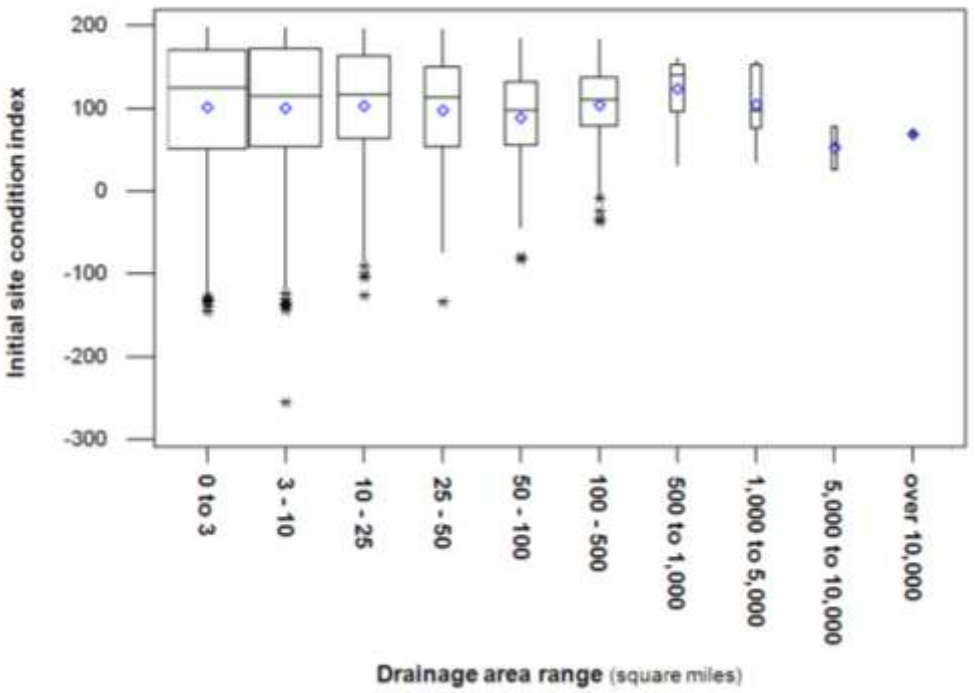


Figure 9. Boxplot of initial site condition index values by drainage area ranges. Diamonds mark the mean index value for each drainage area range. Box widths are proportional to the number of sites in each drainage area range.

The initial site condition index values were divided into bins based on the distribution of values for different sizes of streams (Table 4). Sample sites were grouped into seven groups based on drainage area (i.e., 0-3 square miles, 3-10 square miles, 10-25 square miles, 25-50 square miles, 50-100 square miles, 100-500 square miles, 500-1,000 square miles). Within each of the seven drainage area groups, sites were grouped into six condition tiers based on the initial site condition index based on percentiles of the index distribution in each size group. Sites with initial condition index values greater than the 85th percentile of the index distributions in each size group were designated as "condition 1." Sites with initial condition index values less than the 25th percentile of index distributions in each size group were designated as "condition 6." Sites designated as "condition 2," "condition 3," "condition 4," and "condition 5" were those sites with initial site condition index values between the 85th to 70th, 70th to 55th, 55th to 40th, and 40th to 25th percentiles, respectively. In other words, "condition 1" sites represent sites that were the least impacted by human activities with subsequent tiers representing progressively more impacted sites.

Table 4. Determination of initial site condition categories based on distribution of initial site condition index values. Percentiles were determined for each drainage area grouping.

Drainage area range (square miles)	**Initial Site Condition Category**					
	1	**2**	**3**	**4**	**5**	**6**
0 to 3	**> 85th percentile** of initial site condition index values	**85th to 70th percentile** of initial site condition index values	**70th to 55th percentile** of initial site condition index values	**55th to 40th percentile** of initial site condition index values	**40th to 25th percentile** of initial site condition index values	**< 25th percentile** of initial site condition index values
3 to 10						
10 to 25						
25 to 50						
50 to 100						
100 to 500						
500 to 1,000						

Sites without physical habitat data – about 10% of all sites – were assigned to initial site condition categories of zero (0). Sites from streams draining > 1,000 square miles were not assigned into condition categories since there were only samples from 26 sites draining that much land; each of these samples were evaluated individually.

The percentiles of initial site condition index values chosen as breakpoints between condition categories were selected such that each category would have a reasonably comparable number of samples, and such that "condition 1" and "condition 2" sites would represent the least impacted conditions possible. Other approaches to defining site conditions often set threshold values for each of a suite of abiotic components (e.g., greater than 85% forested land use, less than 5% abandoned mine lands). The condition index approach is analogous to this component-by-component threshold approach since the condition index is built from a suite of abiotic components.

The percentile breakpoints were applied to different groupings of stream sizes because the characteristics of the least disturbed small headwater brooks may be quite different from the least disturbed larger rivers. For example, there are many small streams in Pennsylvania that drain basins with greater than 90% forested land use, but there are relatively few larger rivers that have this high a proportion of forested land in their upstream basins. Applying the site condition index percentile breakpoints to different sizes of streams facilitated distinguishing the least disturbed streams in various size ranges and maintaining stringent standards for what constitutes minimally disturbed streams. This is a key component of such a project since we know and expect that benthic macroinvertebrate communities exhibit natural changes with stream size (Vannotte et al. 1980).

Over 39% of all "condition 1" samples and over 34% of all "condition 2" samples were from streams with EV aquatic life uses. Over 46% of all "condition 1" samples and over 38% of all "condition 2" samples were from streams with HQ aquatic life uses. Across stream sizes, "condition 1" and "condition 2" sites predominantly represented sites in excellent condition with very high percentages of forested land use and optimal total habitat scores (Figure 10) across the state (see Figure 11 below). These two condition tiers (i.e., "condition 1" and "condition 2") represent the reference conditions for subsequent analyses in this project.

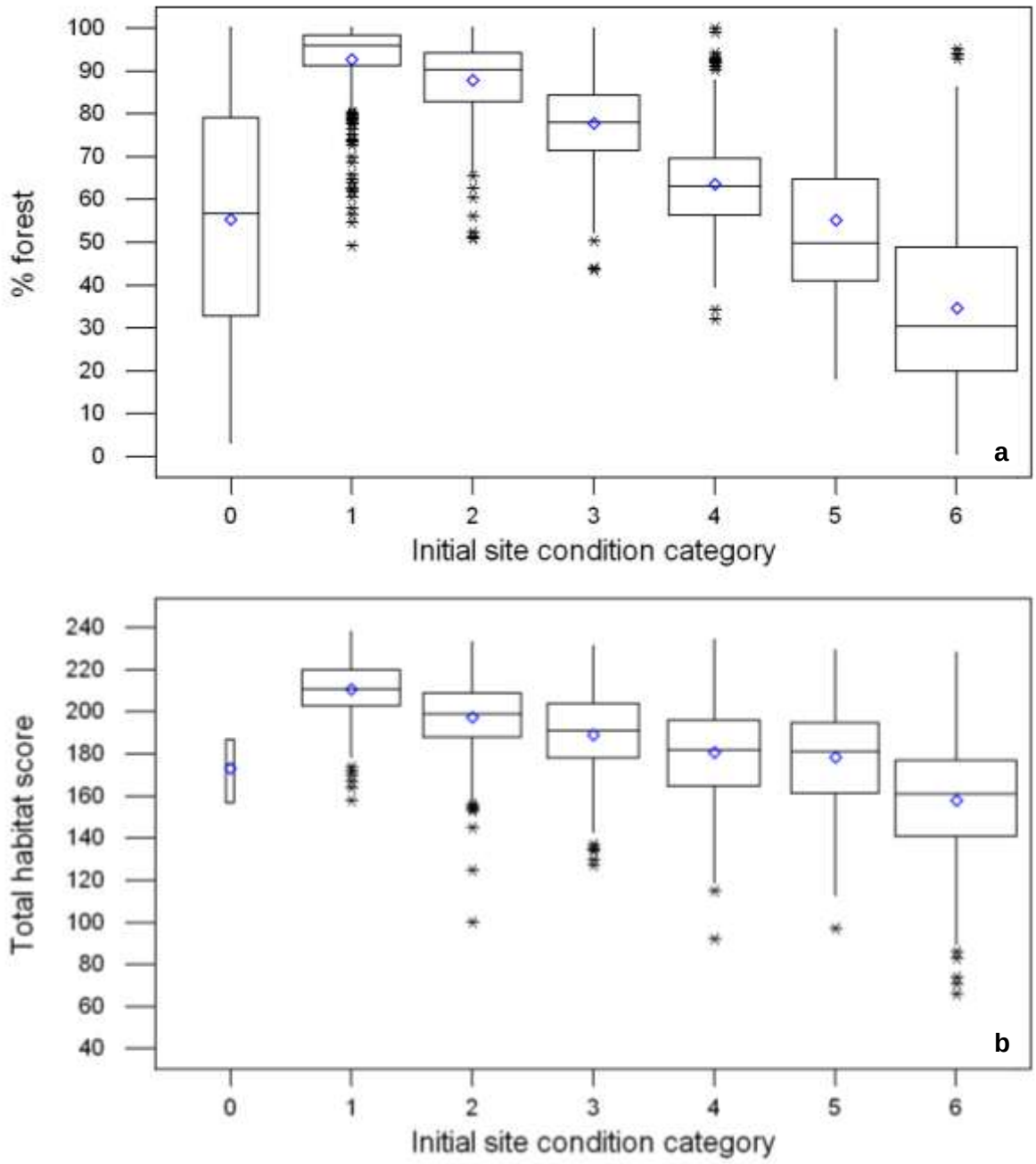


Figure 10. Distribution of (a) percent forested land use in upstream basins and (b) total habitat scores among sites by initial site condition categories.

Field-observed water chemistry data – which consisted of pH, total alkalinity, and specific conductance – were used to further determine site conditions. Any site with a pH recorded below 5.5 was flagged for possible impacts from atmospheric acid deposition, although pH was not recorded for 887 of the 2,482 sites so it is possible that some sites impacted by acidic deposition were not identified. Any site with a specific conductance recorded over 500 µS/cm[c] was considered "condition 6." Specific conductance was not recorded for 837 of the 2,482 sites. The site condition for each sample was assigned as the initial site condition category adjusted for these water chemistry screenings (Table 5).

Table 5. Number of samples by drainage area range and site condition.

Drainage area range (square miles)	Site condition													
	1		2		3		4		5		6		0	
		pH < 5.5		pH < 5.5		pH < 5.5		pH < 5.5		pH < 5.5		pH < 5.5		pH < 5.5
0 to 3	97	33	148	18	118	7	126	3	120	6	244	3	61	1
3 to 10	121	11	108	6	119	5	115	1	96	8	196	5	36	
10 to 25	86	2	76		65		76	1	55	1	118		37	
25 to 50	29		25		30		41		28	3	47		11	
50 to 100	31		26		29		23		19	4	35		4	
100 to 500	38		45		38		29		27		50		29	
500 to 1,000	19		4		3		4		3		16			

DATA EXPLORATION AND SAMPLE CLASSIFICATION

In addition to varying impacts of human activities, natural variation exists among different types of stream ecosystems. For example, biotic assemblages in streams often vary in space and time with basin geology, soil types, stream gradient, substrate composition, climate, and other non-anthropogenic factors. The goal of a classification scheme is to provide a framework for organizing and interpreting the non-anthropogenic spatial and temporal variation of stream ecosystems in order to establish meaningful reference conditions (Whittaker 1962; Hughes 1995; Barbour et al. 1999). Appropriate ecosystem classification is critical to the reference condition concept because it helps determine the spatial and temporal extent to which particular biological attributes apply (Hughes 1995).

Stream classification identifies relatively homogenous classes of streams. Workable classification schemes are characterized by biological expectations that vary less within each class of streams than among the different classes. Representative sites can be selected from each class of streams to establish reference conditions (Barbour et al. 1999). Classification across heterogeneous classes may result in misrepresentation of the biological condition in certain ecosystem types. For these reasons, the need for some sort of classification scheme that groups streams together that are more similar than others (e.g., true limestone spring streams, freestone streams) should be carefully evaluated (Hughes 1995). Evaluation of biological attributes that represent structures and functions of reference condition communities represents a critical component of any classificatory analysis of biological data (Hughes 1995). An analysis of taxa sampled from streams in different areas during different seasons can help identify important classifications for biological expectations (Hughes 1995).

In this project, two multivariate statistical methods – agglomerative hierarchical cluster analysis (Lance and Williams 1967; Milligan 1989) and nonmetric multidimensional scaling, or NMDS (Kruskal and Wish 1978; Ludwig and Reynolds 1988) – were used to explore patterns of variation in the biological data as related to abiotic variables, and to evaluate the biological relevance of various potential classification schemes. Both types of analyses, which have been used in similar applications evaluating biological integrity of stream ecosystems (see Barbour et al. 1995; Hawkins and Norris 2000), were performed using SAS ® 9.1 software. The groups defined by the cluster analysis can be thought of as an *a posteriori* classification scheme based solely on characteristics of the biological community, while the other classification schemes tested were determined *a priori* based on physiochemical, biogeographical, and/or seasonal characteristics (Barbour et al. 1999).

All classification analyses were based on matrices of Bray-Curtis similarity measures (Ludwig and Reynolds 1988) calculated on natural log-transformed proportional abundance of taxa. In order to minimize variation attributable to anthropogenic impacts, all classification analyses were based only on the 923 samples from the 715 reference sites (i.e., “condition 1” or “condition 2’) (Figure 11). These 923 least-disturbed samples contained 293 taxa. Extremely rare taxa (i.e., those encountered in less than five of the 923 samples) were not included in the classification analyses, which resulted in excluding the rarest 101 taxa from the classification analyses and including 192 more common taxa. Previous analyses (see Marchant 1999, 2002) suggest that extremely rare taxa are largely unimportant to multivariate analyses, especially when considering only relatively

undisturbed sites, because only more commonly encountered taxa can be adequately characterized in terms of response to environmental variables. In addition, extremely rare taxa are more likely to have been misidentified and could obscure the ability to detect biologically significant differences among sites (Hawkins et al. 2000).

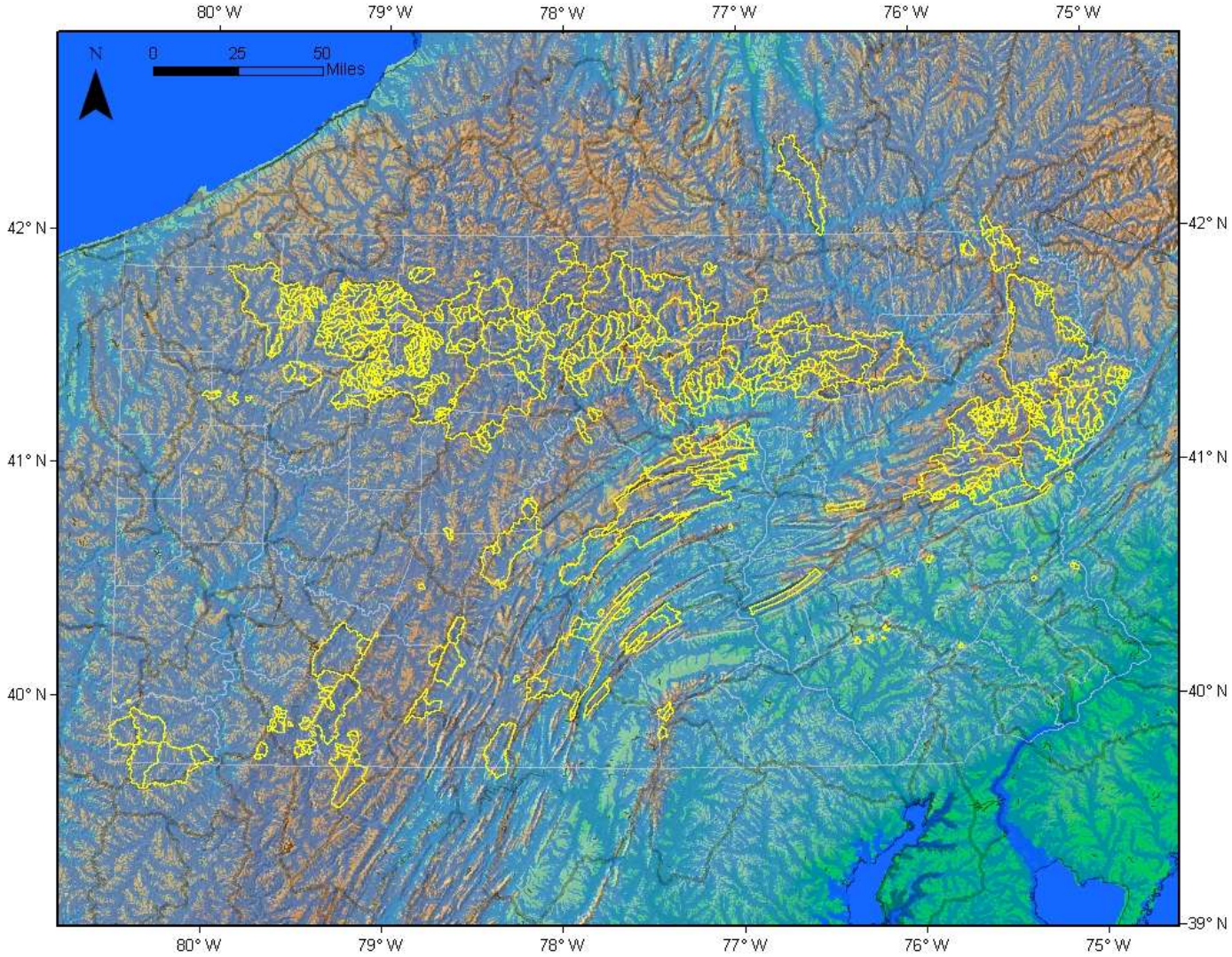


Figure 11. Map of the 715 basins for the 923 samples from reference (i.e., condition 1 and condition 2) sites used in cluster and NMDS analyses.

Cluster Analysis

The cluster tree resulting from the SAS ® CLUSTER procedure using the flexible beta method with a beta value of -0.25 was analyzed at the level of 11 clusters (Figure 12), which explained 27% of the variation in the data. For purposes of the cluster analysis Bray-Curtis similarity measures were converted to distance measures by subtraction from one. The beta value of -0.25 was chosen based on literature (Milligan 1989) and visual inspection of cluster trees constructed using other beta values; a value of -0.25 produced a tree with visually distinguishable groupings, as opposed to other values that tended to produce overly detailed groups (more positive beta values) or overly simplified groups (more negative beta values).

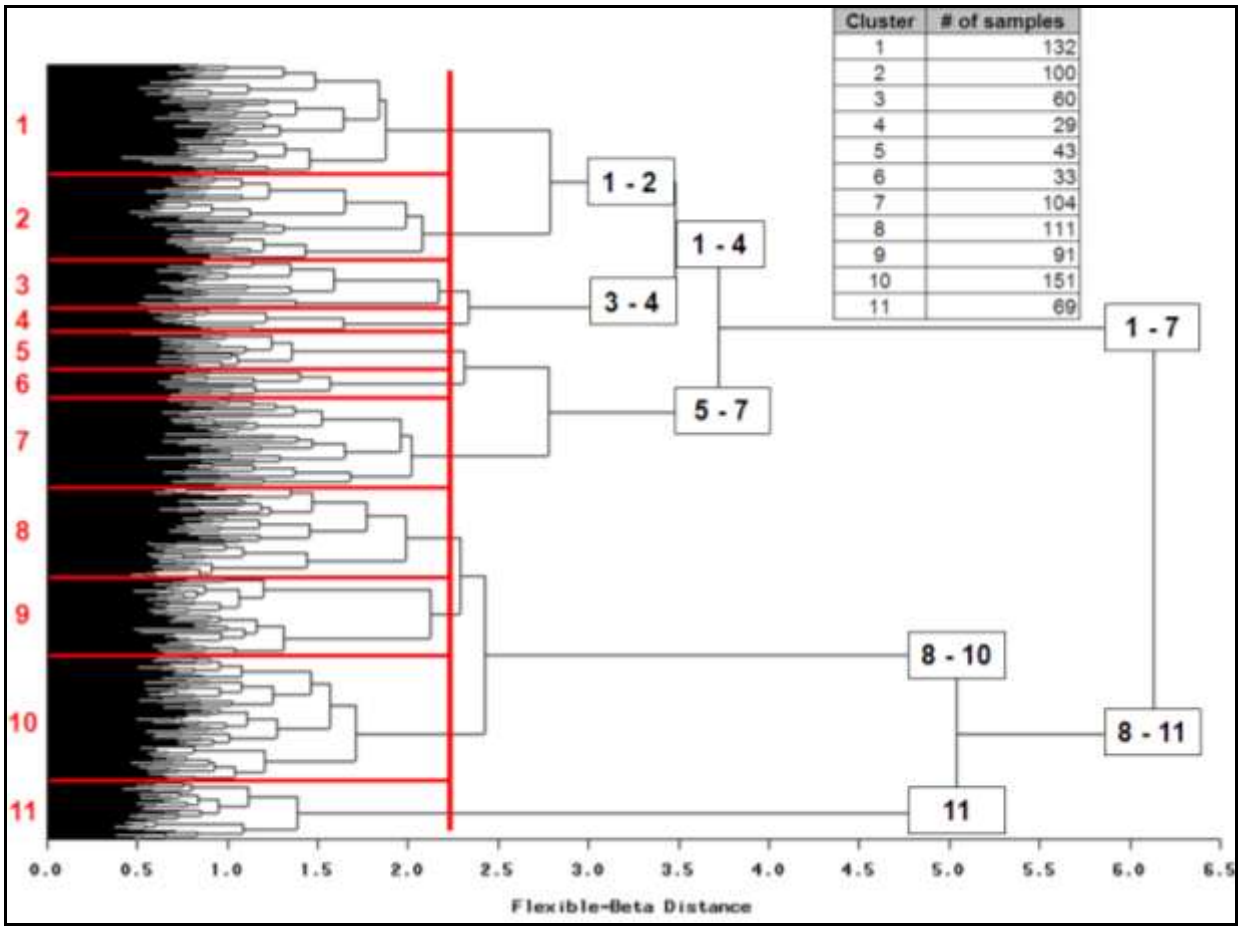


Figure 12. Cluster tree for the 192 most common taxa from 923 reference samples.

At the 11-cluster level, the first break in the cluster tree occurs between clusters 1-7 and clusters 8-11, with the second break between clusters 8-10 and cluster 11, the third break between clusters 1-4 and clusters 5-7, and the fourth break between clusters 1-2 and 3-4.

Condition category

As evidenced by the very even distribution of samples from “condition 1” and “condition 2” sites in each cluster (Table 6), the influence of anthropogenic impacts between these two clusters accounts for very little of the variation in taxa patterns among samples in the cluster analysis. As shown below, much more of the variation in the cluster analysis is accounted for by natural factors such as stream size and sampling season. This provides support for the argument that including samples from both “condition 1” and “condition 2” sites in the cluster analysis does not introduce substantial variation attributable to human impacts.

Table 6. Number of samples in each cluster by condition category.

condition category	Cluster										
	1	2	3	4	5	6	7	8	9	10	11
1	64	46	28	17	19	11	41	43	43	82	
2	67	54	32	12	14	22	47	65	36	64	1
acid	1				10		16	3	12	5	68

Drainage area

Clusters 8-11 contained samples mostly from sites on smaller streams (Figure 13). Of the 422 samples in clusters 8-11, only six samples were from stream sites draining more than 25 square miles of land and no samples were from stream sites draining more than 50 square miles of land. Of the 422 samples in clusters 8-11, 55% drained less than three square miles of land and 90% drained less than ten square miles of land. Samples in cluster 11 were from especially small streams. Of the 69 samples in cluster 11, 60 were from stream sites draining less than three square miles of land and the other nine samples were from stream sites draining between three and ten square miles of land.

Clusters 5 and 7 contained samples mostly from sites on smaller streams, with over 70% of the samples in each of those two clusters being from stream sites draining less than 10 square miles of land and over 90% coming from stream sites draining less than 25 square miles of land. Cluster 6 contained samples mostly from moderate-size and larger stream sites, with over 80% of the 33 samples in that cluster coming from stream sites draining more than 50 square miles of land.

Clusters 1-4 contained samples from sites draining the largest streams in the cluster analysis dataset. Of the 321 samples in clusters 1-4, only eight were from sites on streams draining less than three square miles of land and only 48 were from sites on streams draining less than ten square miles of land. Of all 163 samples in the cluster analysis dataset from streams draining more than 50 square miles of land, 85% were in clusters 1-4. Of the 106 samples from stream sites draining over 100 square miles of land, 67% were in clusters 3 and 4. Samples in clusters 1 and 2 were mostly from more moderate-sized streams, while samples in clusters 3 and 4 were mostly from larger streams.

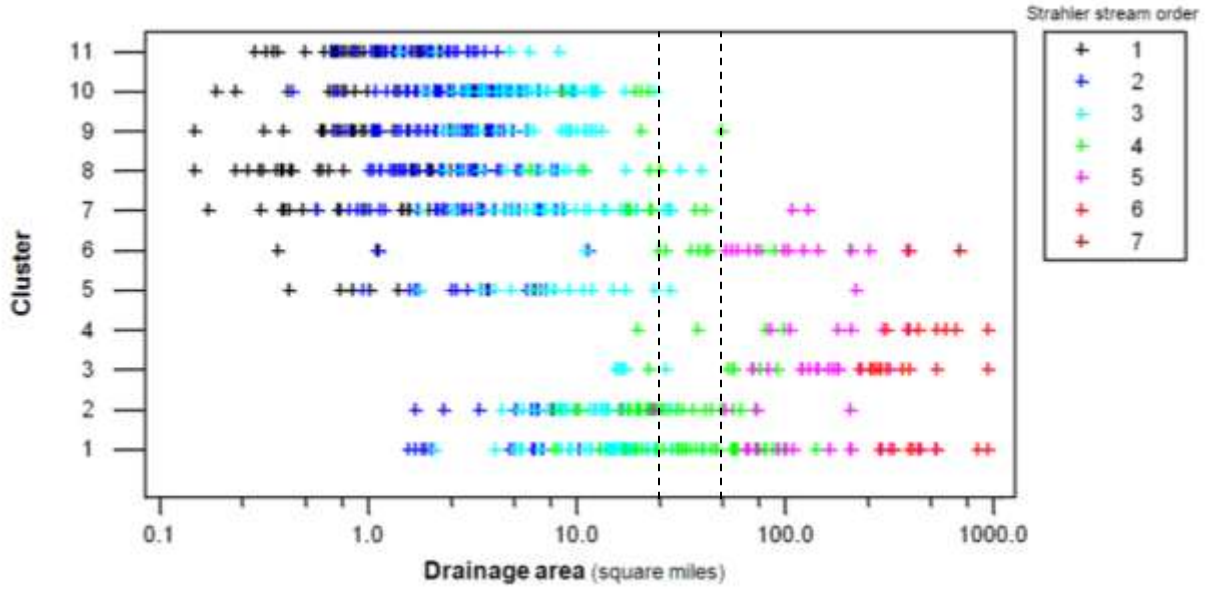


Figure 13. Distribution of sample site drainage areas by cluster, coded by Strahler stream order. Note logarithmic scale for drainage area. Lines are drawn at 25 and 50 square miles.

Stream order

If we look at the cluster results in terms of Strahler stream order we – not surprisingly – see similar patterns as with upstream drainage area (Table 7). Over 95% of the samples in clusters 8-11 were from 1st through 3rd order streams, with no samples from streams larger than 4th order. Cluster 11 contained mostly samples from 1st through 2nd order streams, with no samples from streams larger than 3rd order.

Clusters 5 and 7 were also over 90% comprised of samples from 1st through 3rd order streams, with a few 4th and 5th order samples. About 73% of the samples in cluster 6 were from 4th or 5th order streams, with at least one sample from every stream order represented in this cluster.

Over 85% of the samples in clusters 1 and 2 were from 3rd through 5th order streams, with no samples from 1st order streams. All of the samples in cluster 4 and over 85% of the samples in cluster 3 were from 4th through 6th order streams, with no samples smaller than 4th order in cluster 4 and no samples smaller than 3rd order in cluster 3.

Table 7. Number and percentage of samples in each cluster by Strahler stream order.

Strahler stream order	Cluster																					
	1		2		3		4		5		6		7		8		9		10		11	
	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%
1									7	**16%**	2	**6%**	16	**15%**	31	**28%**	21	**23%**	18	**12%**	29	**42%**
2	12	**9%**	11	**11%**					14	**33%**	2	**6%**	36	**35%**	54	**49%**	39	**43%**	77	**51%**	34	**49%**
3	50	**38%**	43	**43%**	7	**12%**			21	**49%**	2	**6%**	43	**41%**	21	**19%**	28	**31%**	49	**32%**	6	**9%**
4	45	**34%**	38	**38%**	8	**13%**	4	**14%**			11	**33%**	6	**6%**	5	**5%**	3	**3%**	7	**5%**		
5	17	**13%**	8	**8%**	23	**38%**	11	**38%**	1	**2%**	13	**39%**	3	**3%**								
6	8	**6%**			22	**37%**	14	**48%**			2	**6%**										
7											1	**3%**										

Slope

Samples in clusters 8-11 were collected mostly from streams with slopes over 2% (Figure 14). Samples in cluster 7 were also collected mostly form higher gradient streams, with samples in clusters 1, 2, and 5 collected mostly from more moderate gradient (i.e., 0.5% to 2.0%) streams and samples in clusters 3, 4, and 6 collected from mostly lower slope (< 0.5%) streams. These results reflect the inverse relationship of stream slope and drainage area noted earlier (Figure 5).

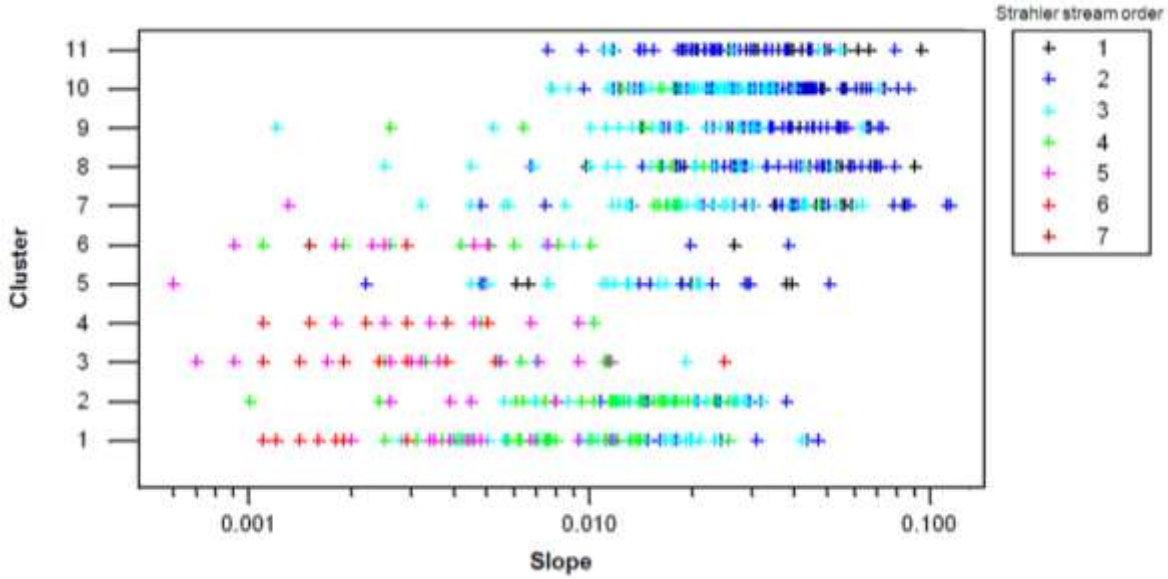


Figure 14. Distribution of sample site slopes by cluster, coded by Strahler stream order. Note logarithmic scale for slope.

Elevation

Patterns of elevation (Figure 15) in the cluster analysis were much less distinct than drainage area, stream order, and slope patterns.

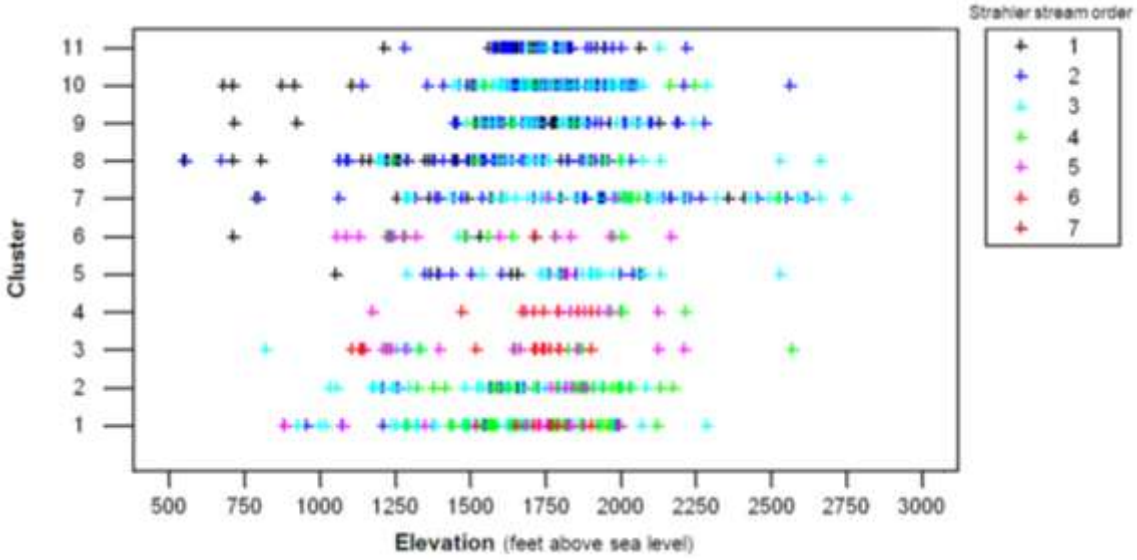


Figure 15. Distribution of sample site elevations by cluster, coded by Strahler stream order.

Seasons

Clusters 8-11 contained samples mostly from the spring months, March through May (Table 8, Figure 16). This was particularly the case for clusters 9-11 where over 90% of the samples in those clusters were March through May samples. In cluster 11, all of the samples were collected March through May.

Samples in cluster 7 were mostly collected late summer, autumn, and early winter, with 81% of samples in this cluster collected August through December. Samples in clusters 5 and 6 were mostly collected late autumn, winter, and early spring, with all the samples in cluster 6 and 93% of the samples in cluster 5 collected November through March.

Samples in cluster 4 were all mostly collected late summer and autumn, with 97% of samples in this cluster collected August through November. Samples in cluster 3 were also concentrated in late summer and early autumn, with 64% of samples in this cluster collected August through October. Samples in cluster 2 were somewhat bimodal in terms of sampling season with 44% of samples in this cluster collected in November or December, and 45% of samples collected in March and April. Samples in cluster 1 were mostly collected late spring, with 86% of samples in this cluster collected in April and May.

Table 8. Number and percentage of samples in each cluster by month.

month	Cluster																					
	1		2		3		4		5		6		7		8		9		10		11	
	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%	#	%
January	2	**2%**	5	**5%**					1	**2%**	5	**15%**			1	**1%**	1	**1%**	2	**1%**		
February	1	**1%**	2	**2%**					1	**2%**	3	**9%**			17	**15%**			9	**6%**		
March	8	**6%**	20	**20%**	2	**3%**			12	**28%**	4	**12%**	3	**3%**	13	**12%**	4	**4%**	30	**20%**	10	**14%**
April	58	**44%**	25	**25%**	8	**13%**			17	**40%**			12	**12%**	38	**34%**	26	**29%**	70	**46%**	33	**48%**
May	56	**42%**	3	**3%**	2	**3%**			3	**7%**			3	**3%**	17	**15%**	58	**64%**	37	**25%**	26	**38%**
June	2	**2%**			3	**5%**									2	**2%**	1	**1%**	1	**1%**		
July	1	**1%**			3	**5%**							2	**2%**			1	**1%**				
August	1	**1%**			10	**17%**	8	**28%**					24	**23%**	1	**1%**						
September	2	**2%**			22	**37%**	10	**34%**					16	**15%**								
October			1	**1%**	6	**10%**	2	**7%**					12	**12%**	2	**2%**						
November			23	**23%**	1	**2%**	8	**28%**	5	**12%**	4	**12%**	21	**20%**	11	**10%**			2	**1%**		
December	1	**1%**	21	**21%**	3	**5%**	1	**3%**	4	**9%**	17	**52%**	11	**11%**	9	**8%**						

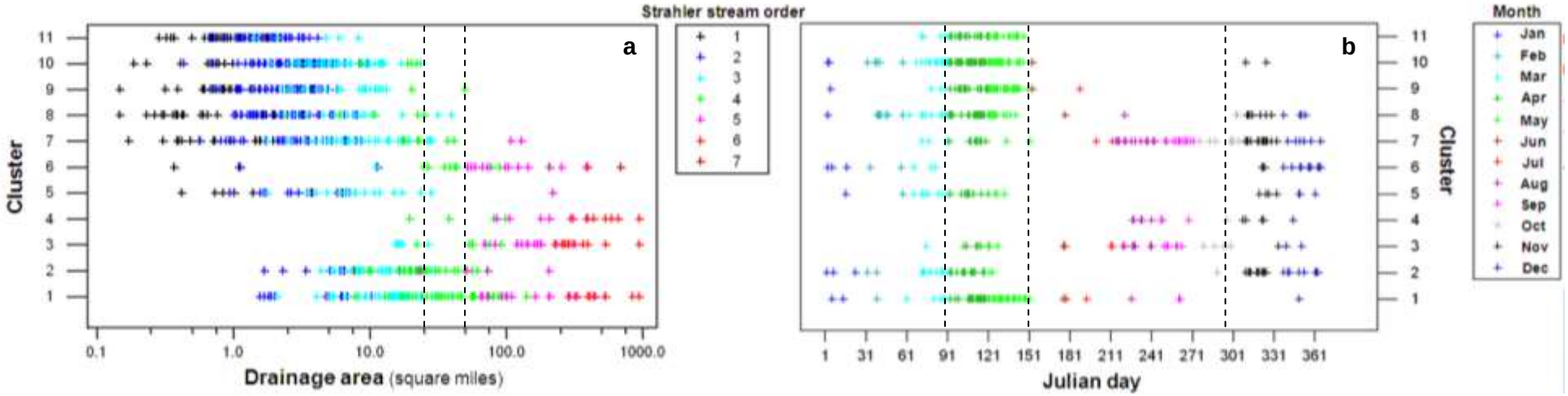


Figure 16. Side-by-side plots of sample site (a) drainage area and (b) Julian day of sample collection in each cluster. Note logarithmic scale for drainage area.

Location

Some patterns in latitude and longitude coordinates of sample locations in each cluster are apparent (Figure 17). However, the patterns in latitude and longitude among clusters were not as consistent among the first, second, third, and fourth major breaks in the cluster tree as were patterns of drainage area and – less so – sampling season. Maps of basin locations for each cluster are presented in Appendix B.

Looking at clusters 8-11, most samples in clusters 9, 10, and 11 were from sites north of the 41st parallel, with a number of sites in more southerly parts of Pennsylvania, while samples in cluster 8 were from sites a touch further south than the samples in cluster 9,10, and 11. Samples in cluster 11 were mostly from western parts of the state, concentrated around the 79th meridian (largely reflecting the density of sites in the Clarion River and Tionesta Creek basins). Samples in cluster 9 were also mostly from sites in the western parts of the state, west of the 78th meridian, while samples in cluster 8 were mostly from sites east of the 78th meridian, and samples in cluster 10 were spread fairly evenly across the state from east to west.

Looking at clusters 5-7, samples in cluster 5 were mostly from sites north of the 41st parallel. Samples in cluster 6 were also from sites mostly north of or near the 41st parallel, but there were also eight samples in cluster 6 from sites south of 40.5° north latitude. Samples in cluster 7 were from sites fairly well distributed in the state from north to south. Samples in clusters 5, 6, and 7 were fairly spread across the state from east to west, although samples in cluster 5 were from sites more concentrated in the eastern parts of the state (particularly the Pocono Plateau), with samples in cluster 6 mostly from sites more in the east-west center of the state (but also including four sites in the very southwest corner of the state in Greene County), and samples in cluster 7 being more from sites in western parts of the state (with a number of samples in the Allegheny Mountains of eastern Fayette County and western Somerset County).

Looking at clusters 3 and 4, there was only one site (on Aughwick Creek) in cluster 4 that was south of the 41st parallel. Samples in cluster 3 were also mostly from sites north of the 41st parallel, but with a number of samples from sites in the southcentral and southwestern parts of the state. Samples in cluster 3 were from sites spread across the state from east to west, while samples in cluster 4 were very concentrated between the 77th and 79th meridian (showing the heavy concentration of sites in this cluster from south-draining tributaries of the West Branch Susquehanna River).

Looking at cluster 1 and 2, the only two samples in cluster 2 from sites south of 40.8° north latitude were from two sites on Dunbar Creek in central Fayette County. Most of the samples in cluster 1 were also from more northerly parts of the state, but this cluster also included many samples from sites in more southerly parts of the state. East to west, samples in clusters 1 and 2 were from sites a bit more in the eastern parts of the state, but with fairly decent east-west spread across the state.

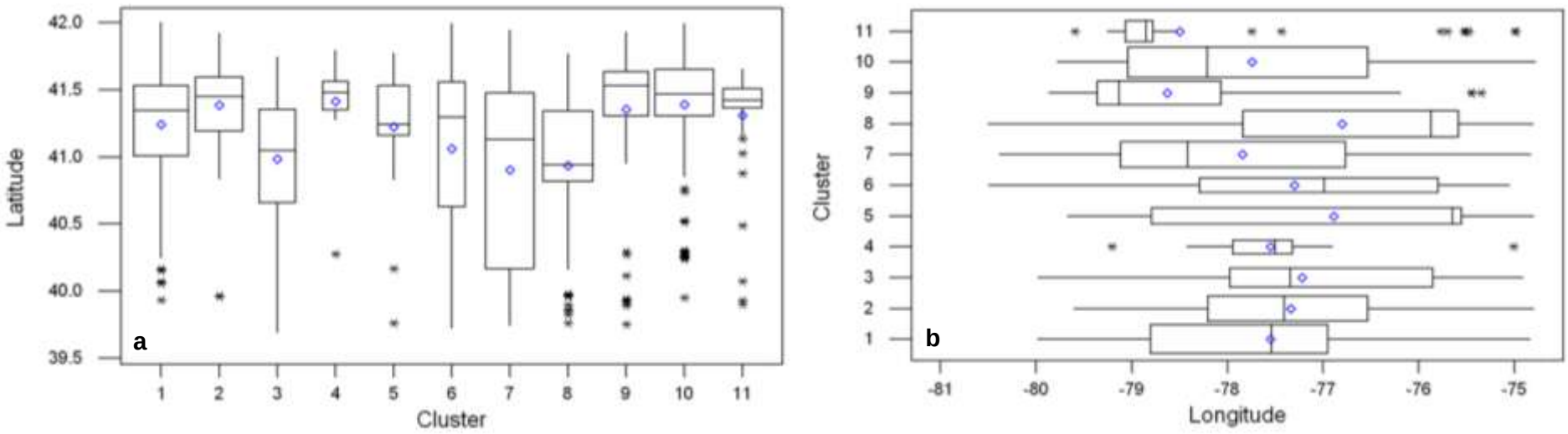


Figure 17. Distribution of sample site (a) latitude and (b) longitude by cluster.

The remainder of the discussion in this section looks at sample locations cluster-by-cluster and basin-by-basin in more detail (Table 9 and see Appendix B). The basins described below are those defined by the United States Geological Survey in their Hydrologic Unit system (Seaber et al. 1987) at the eight-digit level.

Samples in cluster 11 were highly concentrated in northwestern parts of the state, particularly in the upper and middle reaches of the Clarion River basin and in the southern part of the Tionesta Creek basin in eastern Forest County and western Elk County. In fact, 49 of the 69 samples in cluster 11 were located in the Clarion River basin, with another six located in the Tionesta Creek basin. With 71% of the samples in cluster 11 located in the Clarion River basin, this cluster exhibited the highest degree of geographical concentration of any cluster. Cluster 11 also contained six samples from the upper and middle reaches of the Lehigh River basin as well as: two samples from an unnamed tributary to Williams Run in the Conemaugh River basin; one sample from Long Pine Run in the headwaters of the Conococheague Creek basin draining off South Mountain; one sample from Dothan Run – a tributary to Conodoguinet Creek draining off Kittatinny Mountain in northern Franklin County; two samples from an unnamed tributary to Shohola Creek in north-central Pike County; and two samples from Rock Run – a tributary to Dunbar Creek draining the western part of Chestnut Ridge in central Fayette County.

Samples in cluster 10 were particularly concentrated in the Middle Allegheny River - Tionesta Creek basin in southern Warren County and northern Forest County with 44 of the 151 samples (or 43%) in this cluster located in this area. There were also 14 samples from the nearby Clarion River basin in cluster 10. There were also quite a few samples in cluster 10 located in the Lower West Branch Susquehanna River basin, particularly in the upper reaches of Loyalsock Creek in Sullivan County and eastern Lycoming County. Cluster 10 also contained samples from sites in other parts of the state.

Samples in cluster 9 were also primarily concentrated in the Middle Allegheny River - Tionesta Creek basin and secondarily in the neighboring Clarion River basin, but less heavily than clusters 11 and 10. Other samples in cluster 9 were located in various parts of the Commonwealth.

Samples in cluster 8 were noticeably less concentrated in the vicinities of the Tionesta Creek and Clarion River basins. In fact, the basin with the most samples in cluster 8 was the Lehigh River basin, where 36 of the 111 samples in the cluster were located. There were a few handfuls of samples in cluster 8 located in other parts of the upper and middle Delaware River drainages as well some samples from other regions of Pennsylvania.

Samples in cluster 7 were more distributed around the state. The basin with the most samples in cluster 7 was the Youghiogheny River basin with 22 of 104 samples located in the highlands of this basin in eastern Fayette County and western Somerset County. The three samples from sites draining over 100 square miles of land in cluster 7 were located on West Branch Tionesta Creek and East Branch Clarion River.

Samples in cluster 6 were also distributed around Pennsylvania, with many samples located in the West Branch Susquehanna River basin. Cluster 6 contained three samples from White Deer Hole Creek, two samples from the lower reaches of Loyalsock Creek, one sample from the lower reaches of Lycoming Creek, and one sample from White Deer

Creek, all in the Lower West Branch Susquehanna River basin. This cluster also contained three samples from Brodhead Creek, two samples from Dunkard Creek (and one from Dunkard Fork), two samples from Sherman Creek as well as samples from streams in other parts of the state.

Samples in cluster 5 were primarily concentrated in the upper Lehigh River basin and secondarily in the Middle Delaware River basin as well as the Tionesta Creek and Clarion River basins with a few samples from other parts of the state. The only sample in cluster 5 from a site draining more than 30 square miles of land was from a site on Kettle Creek draining about 220 square miles.

Samples in cluster 4 were concentrated in the northcentral part of the state. Twelve of the 29 samples in cluster 4 were from the main stem of Pine Creek in eastern Potter County, southwestern Tioga County, and northwestern Lycoming County at sites ranging in drainage area from 38 square miles to 940 square miles. This cluster also included three samples from different sites along First Fork Sinnemahoning Creek, two samples from a site on Driftwood Branch Sinnemahoning Creek, two samples from different sites along Kettle Creek, two samples from a site on Lycoming Creek, three samples from two sites along lower reaches of Loyalsock Creek, and two samples from a site on Aughwick Creek as well as one sample each from the Lackawaxen River, Potato Creek, and Tionesta Creek. The only sample in this cluster from a site draining less than 35 square miles was a sample from the upper reaches of First Fork Sinnemahoning Creek.

Samples in cluster 3 were also concentrated in the northcentral part of the state with nine samples from the middle and lower reaches of Pine Creek in eastern Potter County, southwestern Tioga County, and northwestern Lycoming County as well as one sample from near the mouth of Little Pine Creek. There were also five samples from Spruce Run in Union County in cluster 3 as well as two samples from the lower reaches of Loyalsock Creek and one sample from a site on Muncy Creek. Cluster 3 also contained seven samples from the lower reaches of Brodhead Creek as well as a sample from two other nearby creeks: Marshalls Creek and Shohola Creek. Eight samples from Aughwick Creek also clustered into cluster 3 as well as some other samples from other parts of the state. Five of the six samples from sites draining less than 25 square miles of land in cluster three were from Spruce Creek in Union County, with the other one from Trout Creek in Monroe County.

Samples in cluster 2 were most heavily concentrated in the northcentral part of the state. This cluster contained 25 samples from the Lower West Branch Susquehanna River basin, with six samples from Muncy Creek, three samples from White Deer Hole Creek, and samples from many other streams in this area. There were quite a few samples from the Middle Allegheny River - Tionesta Creek basin in cluster 2, with six samples from West Branch Caldwell Creek, four samples from East Hickory Creek, and a number of other samples from other creeks in this area. Nine samples in cluster 2 were located at various sites along Sinnemahoning Portage Creek with two other samples from a site in the lower reaches of First Fork Sinnemahoning Creek. This cluster also contained many samples from the Middle Delaware River basins and other samples from around the state. The only samples from sites draining over 75 square miles in cluster 2 were from a site on First Fork Sinnemahoning Creek draining 205 square miles. The next largest site in this cluster was a site draining 73 square miles on Sinnemahoning Portage Creek.

Samples in cluster 1 were spread about the state quite a bit. There were over 15 samples in this cluster from four different basins: 22 from a variety of streams in the Middle Allegheny River - Tionesta Creek basin; 20 from various streams in the Lower West Branch Susquehanna River basin; 17 from the Middle West Branch Susquehanna River basin (11 of those from Kettle Creek); and 16 samples from different streams in the Lehigh River basin. Cluster 1 also included samples from many other parts of the state.

Table 9. Number of samples in each cluster by basin. Basins are those defined by the United States Geological Survey's Hydrologic Unit system (Seaber et al. 1987) at the eight-digit level.

Hydrologic Unit		Cluster										
Code	Name	1	2	3	4	5	6	7	8	9	10	11
02040101	Upper Delaware		1				2		1			
02040103	Lackawaxen	5	6	1	1	2			4	1	9	
02040104	Middle Delaware-Mongaup-Brodhead	4	11	9		6	3	12	15		4	2
02040105	Middle Delaware-Musconetcong						1	1	3			
02040106	Lehigh	16		5		15		3	36		7	6
02040203	Schuylkill							3	1		3	
02050101	Upper Susquehanna	2					2	1			5	
02050103	Owego-Wappasening	1										
02050104	Tioga							1		5	3	
02050106	Upper Susquehanna-Tunkhannock		2				2	3		2		
02050107	Upper Susquehanna-Lackawanna		1			2		2		3	2	
02050201	Upper West Branch Susquehanna							8		1	2	
02050202	Sinnemahoning	2	13	3	5		1	3		1	11	
02050203	Middle West Branch Susquehanna	17	11	1	2	1	1	2	2	1	5	
02050204	Bald Eagle	8	4								2	
02050205	Pine	9	7	10	12	1	1	3	3	3	4	
02050206	Lower West Branch Susquehanna	20	25	9	5		8	7	9	5	21	
02050301	Lower Susquehanna-Penns			1				1	2		1	
02050302	Upper Juniata			4				3	1		1	
02050303	Raystown	2					1			1	2	
02050304	Lower Juniata	5		8	2		1	1	1			
02050305	Lower Susquehanna-Swatara	4					2	1				1
02050306	Lower Susquehanna								4	2	4	
02070003	Cacapon-Town			1								
02070004	Conococheague-Opequon	1							2			1
05010001	Upper Allegheny	4	1		1		2	5	1		3	
05010003	Middle Allegheny-Tionesta	22	16	4	1	6	2	6	10	39	44	6
05010004	French										1	
05010005	Clarion	7		1		7		11	3	20	14	49
05010007	Conemaugh					1		4	1		1	2
05010008	Kiskiminetas	2									1	
05020004	Cheat					2			1	1		
05020005	Lower Monongahela			1			3		2			
05020006	Youghiogheny	1	2	2				22	4	6	1	2
05030105	Connoquenessing								1			
05030106	Upper Ohio-Wheeling						1	1	4			

Note that different samples from the same site could – and often did – appear in different clusters. For example, one site in the lower reaches of Loyalsock Creek in Lycoming County near Butternut Grove was sampled eight different times and these eight samples ended up in four different clusters: one in cluster 1; three in cluster 3; two in cluster 4; and two in cluster 6. The differential clustering of samples from the same site in this instance appears directly related to sampling season: the one sample in cluster 1 was sampled in mid-April; the three samples in cluster 3 were sampled June, July, and August; the two samples in cluster4 were sampled mid-October and early December; and the two samples in cluster 6 were sampled late December and mid-January. Of all the sites in the cluster analysis dataset, 58 had samples that ended up in two different clusters, 10 had samples that ended up in three different clusters, and the one site on Loyalsock Creek had samples that ended up in four different clusters.

Taxa

The biotic characteristics of each cluster were thoroughly explored by tabulating abundance (i.e., number of individuals per sample in each cluster) and occurrence (i.e., percent of samples encountered in each cluster) of each taxon in each cluster. The abundance and occurrence of each taxon was ranked for each cluster. Taxa were sorted by abundance rank and abundance to determine which taxa were uniquely abundant – and uniquely scarce or absent – in each cluster. Taxa were then sorted by occurrence rank and occurrence to determine which taxa were uniquely encountered most often – and least often – in each cluster. Taxa were sorted by rank first to determine patterns unique to each cluster. As an example, Chironomidae were fairly abundant and nearly ubiquitously encountered in every cluster. Ranking the abundance and occurrence provides a picture of how those patterns vary cluster to cluster in relative terms. Only the 192 taxa included in the cluster analysis were considered in this tabulation. Rather than go through detailed taxa patterns cluster by cluster – which would require substantial space – the following discussion focuses on the patterns on taxa abundance and occurrence related to the major breaks in the agglomerative cluster tree. In other words, the following analysis focuses on the differences in taxa driving the major branches in the cluster tree. More details on this analysis are available upon request.

A few taxa exhibited distinct abundance and/or occurrence patterns at the first break in the cluster tree – between clusters 1-7 and clusters 8-11 (Table 10). The following taxa were much more commonly encountered in clusters 1-7 than clusters 8-11 (taxa with the largest differences in abundance and occurrence between these two cluster groups are listed first): Cheumatopsyche and Chimarra caddisflies; Ophiogomphus and Stylogomphus dragonflies; Atherix true flies; Isonychia mayflies; Ancylidae snails; Ephemera mayflies; Optioservus beetles; Paragnetina stoneflies; Psephenus and Stenelmis beetles; Eurylophella mayflies; Hydracarina water mites; Taeniopteryx stoneflies; Clinocera true flies; Dubiraphia beetles; Allocapnia stoneflies; Ceratopsyche and Glossosoma caddisflies; and Stenonema mayflies. The following taxa were much more commonly encountered in clusters 8-11 than clusters 1-7: Amphinemura stoneflies; Diplectrona caddisflies; Haploperla stoneflies; Probezzia true flies; Leuctra stoneflies; Wormaldia caddisflies; Ameletus mayflies; Hexatoma true flies; Oulimnius beetles; Chelifera true flies; Cambaridae crayfish; Dicranota true flies; Rhyacophila caddisflies; Tallaperla and Alloperla stoneflies; Pteronarcys stoneflies; Polycentropus caddisflies; and Sweltsa stoneflies. Similar patterns were observed for abundance for almost all of the previously mentioned taxa at the first split in the cluster tree.

Table 10. Summary of major taxonomic patterns at the first cluster tree split.

Taxonomic group	Taxa more common and/or abundant in clusters 1-7	Taxa more common and/or abundant in clusters 8-11
Mayflies	Isonychia	Ameletus
	Ephemera	
	Eurylophella	
	Stenonema	
Odonates	Ophiogomphus	
	Stylogomphus	
Stoneflies	Paragnetina	Amphinemura
	Taeniopteryx	Haploperla
	Allocapnia	Leuctra
		Tallaperla
		Alloperla
		Pteronarcys
		Sweltsa
Caddisflies	Cheumatopsyche	Diplectrona
	Chimarra	Wormaldia
	Ceratopsyche	Rhyacophila
	Glossosoma	Polycentropus
Beetles	Optioservus	Oulimnius
	Psephenus	
	Stenelmis	
	Dubiraphia	
True Flies	Clinocera	Probezzia
	Atherix	Hexatoma
		Chelifera
		Dicranota
Other Taxa	Ancylidae	Cambaridae
	Hydracarina	

At the second split in the cluster tree, which broke cluster 11 from clusters 8-10, many mayfly taxa were much more commonly encountered in clusters 8-10 than cluster 11, primarily: Cinygmula; Habrophlebiodes; Epeorus; Paraleptophlebia; Acerpenna; Diphetor; Baetis; Ephemerella; and Drunella (Table 11). A few other taxa were also much more common in clusters 8-10 than cluster 11, including: Pteronarcys stoneflies; Ectopria beetles; and Isoperla stoneflies.

In fact, the outstanding patterns in cluster 11 taxa were very high abundance of Leuctra and Amphinemura stoneflies relative to other clusters combined with very low occurrence and abundance of mayflies relative to other clusters. Average Leuctra and Amphinemura abundance per sample in cluster 11 – 65 and 42, respectively – were both at least double that of the next most abundant cluster for those two taxa – cluster 9 average abundances per sample were 27 and 21, respectively. Prosimulium black flies were also relatively abundant in cluster 11 – averaging 25 individuals per sample – compared to clusters 8-10. Mayfly abundance and occurrence was very low in cluster 11 relative to other clusters. The most abundant mayflies in cluster 11 were Eurylophella and Ameletus, but both averaged less than 1 individual per sample. Eurylophella was the most commonly encountered mayfly in cluster 11 samples, but was only found in 22% of all samples in that cluster, ranking seventh among all clusters for occurrence of that taxon.

Table 11. Summary of major taxonomic patterns at the second cluster tree split.

Taxonomic group	Taxa more common and/or abundant in clusters 8-10	Taxa more common and/or abundant in cluster 11
Mayflies	Cinygmula	
	Habrophlebiodes	
	Epeorus	
	Paraleptophlebia	
	Acerpenna	
	Diphetor	
	Baetis	
	Ephemerella	
	Drunella	
Stoneflies	Pteronarcys	Leuctra
	Isoperla	Amphinemura
Beetles	Ectopria	
True Flies		Prosimulium

The third split in the cluster tree broke out clusters 1-4 from clusters 5-7. Taxa that were more commonly encountered in clusters 1-4 than clusters 5-7 (Table 12) included: Acroneuria stoneflies; Baetis, Acentrella, Ephemerella, and Leucrocuta mayflies; Agnetina stoneflies; Serratella and Plauditus mayflies; Optioservus beetles; and Paragnetina stoneflies. Taxa more commonly found in clusters 5-7 than clusters 1-4 included: Eurylophella mayflies; Pycnopsyche caddisflies and Sialis alderflies. Similar abundance patterns were observed for many of these same taxa at the third split in the cluster tree. Allocapnia stoneflies were more abundant in clusters 5-7 – particularly cluster 6 and less so cluster 7 – than in clusters 1-4.

Table 12. Summary of major taxonomic patterns at the third cluster tree split.

Taxonomic group	Taxa more common and/or abundant in clusters 1-4	Taxa more common and/or abundant in clusters 5-7
Mayflies	Baetis	Eurylophella
	Acentrella	
	Ephemerella	
	Leucrocuta	
	Serratella	
	Plauditus	
Stoneflies	Acroneuria	Allocapnia
	Agnetina	
	Paragnetina	
Caddisflies		Pycnopsyche
Beetles	Optioservus	
Other Taxa		Sialis

The fourth split in the cluster tree separated clusters 1-2 and clusters 3-4. Taxa more commonly encountered in clusters 1-2 than clusters 3-4 (Table 13) included: Isoperla stoneflies; Neophylax caddisflies; Sweltsa stoneflies; Probezzia, Prosimulium, and Dicranota true flies; Amphinemura stoneflies; Rhyacophila caddisflies; Leuctra stoneflies; Diplectrona caddisflies; Tallaperla stoneflies; Polycentropus caddisflies; Hexatoma true flies; Haploperla stoneflies; Lanthus dragonflies; Paraleptophlebia mayflies; Cinygmula mayflies; Chelifera true flies; Alloperla stoneflies; Dolophilodes caddisflies; Ectopria beetles;

Pteronarcys stoneflies; as well as Epeorus, Drunella, and Ephemerella mayflies. Taxa more commonly encountered in clusters 3-4 than clusters 1-2 included: Stenelmis beetles; Brachycentrus and Chimarra caddisflies; Heterocloeon, Caenis, and Plauditus mayflies; Cheumatopsyche caddisflies; Argia damselflies; Corydalus dobsonflies; Isonychia and Tricorythodes mayflies; Macrostemmum caddisflies; Paragnetina stoneflies; Leucrocuta and Serratella mayflies; Optioservus beetles; and Acroneuria stoneflies.

Table 13. Summary of major taxonomic patterns at the fourth cluster tree split.

Taxonomic group	Taxa more common and/or abundant in clusters 1-2	Taxa more common and/or abundant in clusters 3-4
Mayflies	Paraleptophlebia	Heterocloeon
	Cinygmula	Caenis
	Epeorus	Plauditus
	Drunella	Isonychia
	Ephemerella	Tricorythodes
		Leucrocuta
		Serratella
Odonates	Lanthus	Argia
Stoneflies	Isoperla	Paragnetina
	Sweltsa	Acroneuria
	Amphinemura	
	Leuctra	
	Tallaperla	
	Haploperla	
	Alloperla	
	Pteronarcys	
Caddisflies	Neophylax	Brachycentrus
	Rhyacophila	Chimarra
	Diplectrona	Cheumatopsyche
	Polycentropus	Macrostemmum
	Dolophilodes	
Beetles	Ectopria	Stenelmis
		Optioservus
True Flies	Probezzia	
	Prosimulium	
	Dicranota	
	Hexatoma	
	Chelifera	
Other Taxa		Corydalus

Discussion

Patterns of occurrence and abundance of various taxa were apparent in the cluster analysis when viewed at the 11-cluster level. These taxonomic patterns correlated primarily to patterns in stream size and sampling season.

The first break in the cluster tree differentiated clusters 1-7 from clusters 8-11. Clusters 8-11 consisted mostly of samples from small, first- through third-order streams draining less than 25 square miles of land. Most of the samples in clusters 8-11 were collected in the

spring months of March, April, and May with a few samples from other times of the year, but very few samples collected June through October. Samples in clusters 8-11, relative to samples in clusters 1-7, were characterized by higher abundance and/or occurrence of Ameletus mayflies, many stonefly genera from a number of families, a handful of caddisfly genera (two of which are net-spinning genera that construct relatively coarse-mesh nets, one of which is a free living genera, and one of which makes a silk tube retreat), Oulimnius beetles, a handful of true fly genera, and a crayfish family.

Clusters 1-7 – as a group – contained samples from larger streams than clusters 8-11. However, clusters 5 and 7 contained samples from streams similar in size to clusters 8-11. Samples in clusters 1-7 represented a range of sampling seasons, with at least a few samples collected during every month of the year. Samples in clusters 1-7, relative to samples in clusters 8-11, were characterized by higher abundance and/or occurrence of a handful of mayfly genera, one stonefly genera in the Perlidae family, two winter stonefly genera, a handful of caddisfly genera (three of which are net-spinning genera that construct relatively fine-mesh nets and one of which typically builds cases out of small pieces of rock), a handful of beetle genera, a couple true fly genera, Ancylidae snails, and water mites.

The data shows that this first split in the cluster tree breaks samples from relatively small streams sampled mostly in the spring (clusters 8-11) distinct from samples from larger streams and smaller streams sampled outside of spring. Samples in clusters 5 and 7 were from streams of similar size as samples in clusters 8-11, but were grouped differently in the cluster tree's first break. Although cluster 7 drained mostly first through third order streams, samples in this cluster were collected mostly during late summer (i.e., August) through early winter (i.e., December). Taxa that were most notably higher in abundance and/or occurrence in cluster 7 than clusters 8-11 include: Cheumatopsyche caddisflies; Taeniopteryx stoneflies; Glossosoma caddisflies; Ephemera and Eurylophella mayflies; Chimarra caddisflies; Isonychia mayflies; and Ceratopsyche caddisflies. Taxa that were most notably higher in abundance and/or occurrence in clusters 8-11 than cluster 7 included: Haploperla and Amphinemura stoneflies; and Diplectrona caddisflies. The difference in abundance and occurrence of Amphinemura stoneflies was particularly dramatic between cluster 7 and clusters 8-11 and was a big driver of the split of cluster 7 from clusters 8-11. Amphinemura stoneflies occurred in at least 75% of all samples in clusters 8-11, while that genus was only seen in 25% of samples in cluster 7. Likewise, Amphinemura stoneflies averaged over 20 individuals per sample in clusters 8-11 and only 2 individuals per sample in cluster 7. Amphinemura stoneflies exhibit one of the most pronounced seasonal booms of any benthic macroinvertebrate taxon in Pennsylvania streams. Larvae of this stonefly genus greatly increase in abundance in small, cold Pennsylvania streams from late March through May. Thus, it is not surprising this taxa was the major driver of a split between samples from small streams sampled in the spring (clusters 8-11) from larger streams (clusters 1, 2, 3, 4, and 6) and small streams sampled at other times of the year (cluster 7).

But what about cluster 5? Samples in cluster 5 were almost exclusively from first through third order streams and were sampled mostly in March and April with a fair percentage of samples from May, November, and December as well. Why did cluster 5 separate from its small stream, springtime comrades in clusters 8-11? Taxa more commonly encountered and/or abundant in cluster 5 than clusters 8-11 included: Promoresia beetles; Eurylophella mayflies; Hydropsyche caddisflies; Prosimulium black flies; Chimarra caddisflies;

Acerpenna mayflies; Sialis alderflies; Paracapnia and Prostoia stoneflies; Stenelmis beetles; Atherix true flies; and Allocapnia stoneflies. Taxa that were more commonly encountered and/or abundant in clusters 8-11 than cluster 5 included: several stonefly genera (Sweltsa, Amphinemura, Pteronarcys, Isoperla, Haploperla, Peltoperla, Alloperla) and mayfly genera (Cinygmula, Ephemerella, Baetis, Epeorus, Drunella, Paraleptophlebia); as well as Hexatoma crane flies; Diplectrona caddisflies; Probezzia blackflies; and Wormaldia caddisflies. However, the biggest drivers of cluster 5 separating from clusters 8-11 appeared to have to do with patterns of taxa abundance. Prosimulium blackflies and Chironomidae midges exhibited much higher abundances in cluster 5 than clusters 8-11. Prosimulium blackflies averaged 58 individuals per sample in cluster 5 and only 15 individuals per sample in clusters 8-11. Chironomidae midges averaged 50 individuals per sample in cluster 5 and only 31 individuals per sample in clusters 8-11. Thus, it appears samples in cluster 5 separated from their small stream, springtime comrades in clusters 8-11 because samples in cluster 5 were dominated by Prosimulium blackflies and less so Chironomidae midges. Prosimulium blackflies are another taxon that exhibits very pronounced seasonal population booms, particularly from December through May. The benthic larvae of Prosimulium blackflies and Chironomidae midges are often very small in size relative to other benthic macroinvertebrates. Individuals of these two taxa often are found in very dense colonies, so that if a particular area containing one of these dense colonies is sampled with one or a few kicks, many hundreds or even thousands of individuals can wash into the net. This can lead to an overwhelming of the sub-sample by Prosimulium blackflies and/or Chironomidae midges, which is what we see in many of the samples in cluster 5. This issue of samples and sub-samples being dominated by large numbers of individuals from one or a few taxa is discussed further below.

The second break in the cluster tree separated cluster 11 from clusters 8-10. This second break does not appear to be driven by differences in stream size or sampling season – almost all samples in clusters 8-11 were from small streams sampled in the spring. The overwhelming pattern in taxa driving the split of cluster 11 from clusters 8-10 – as discussed above – is that samples in cluster 11 were often much more dominated by Leuctra and Amphinemura stoneflies – and less so Prosimulium blackflies and Chironomidae midges than samples in clusters 8-10. Samples in cluster 11 also exhibited marked absence or scarcity of mayflies compared to samples in cluster 8-10. In addition , samples in cluster 11 exhibited the highest degree of geographic concentration of any cluster, with over 75% of samples in cluster 11 located in the upper and middle reaches of the Clarion River basin and in the southern part of the Tionesta Creek basin in eastern Forest County and western Elk County. This part of Pennsylvania has historically received some of the most severe acidic deposition anywhere in the state (Figure 18; see Lynch et al. 2007). In addition, the geologies of the plateau between the Tionesta Creek basin and the Clarion River basin in this area are largely of the Pottsville Formation, which is a Pennsylvanian-age formation primarily made up of sandstone, secondarily conglomerate, with tertiary shale, siltstone, claystone, limestone, and coal. These geologies do not confer much acid buffering capacity to streams draining this landscape. Additionally, Ciolkosz and Levine (1983) identify the soils in this part of the state as being very sensitive to acidic deposition. The severe acid deposition that has impacted this region for many decades, combined with the low geological and edaphic buffering capacity, results in streams that can experience very low pH levels seasonally – especially in early spring – or chronically.

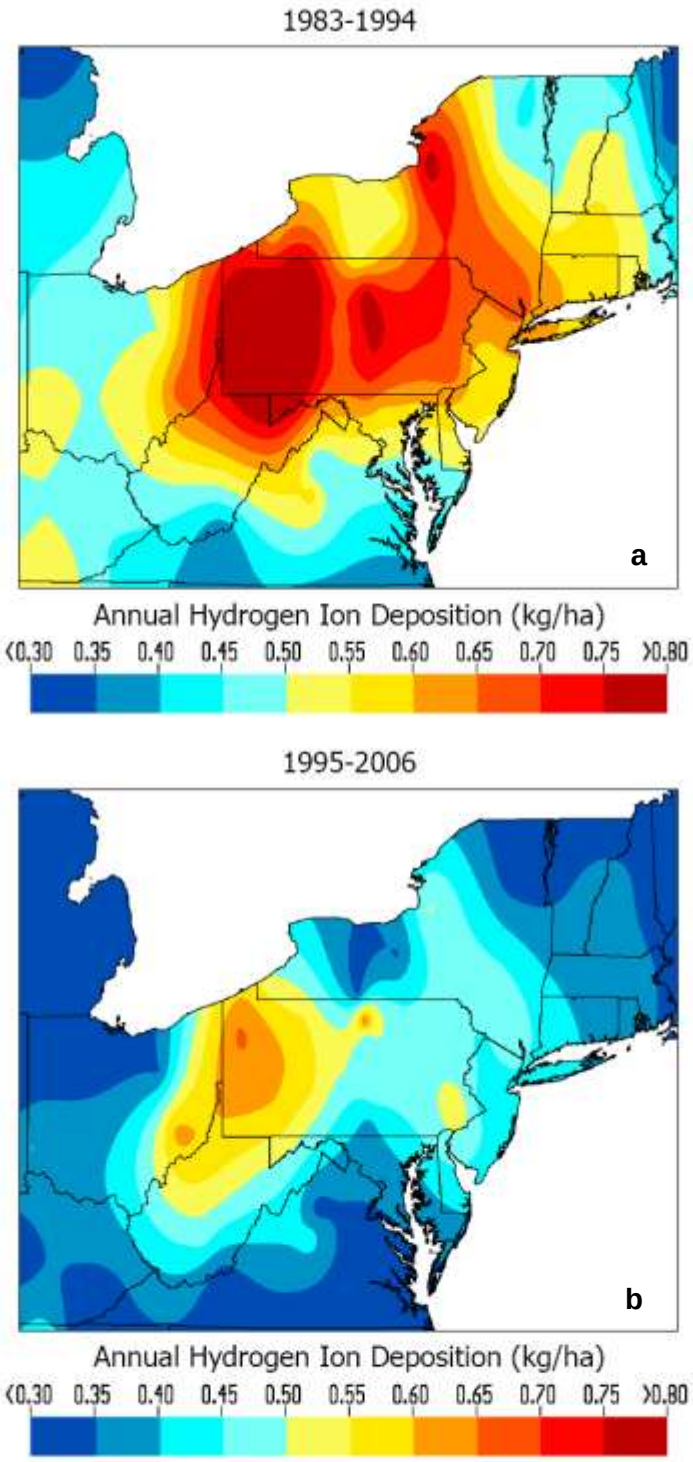


Figure 18. Reproduced from Lynch et al. (2007). Mean annual hydrogen ion deposition across Pennsylvania and neighboring states (a) before (1983-1994) and (b) after (1995-2006) implementation of Title IV of the Clean Air Act Amendments of 1990.

Many studies have documented sensitivity of a number of mayflies and other taxa to acidic conditions in streams (Madarish and Kimmel 2000; Guerold et al. 2000; Kimmel 1999; Griffith et al. 1995; Rosemond et al. 1992; Giberson and Mackay 1991; Simpson et al. 1985; see Sutcliffe and Hildrew 1989) as well as relative tolerance of acidic conditions in streams by some stonefly taxa, particularly Leuctra and Amphinemura (Madarish and Kimmel 2000; Guerold et al. 2000; Kimmel 1999; Griffith et al. 1995; Griffith et al. 1994; Simpson et al. 1985). Taken together, these lines of evidence suggest that the samples in cluster 11 reflect streams impacted by acid deposition. More consideration to acid deposition is presented below.

The third break in the cluster tree separated clusters 1-4 from clusters 5-7. Characteristics of samples in clusters 5 and 7 were already discussed. Samples in cluster 6 were mostly from fourth and fifth order streams, but streams sites in this cluster ranged from first order through seventh order. Samples in cluster 6 were all collected November through March. The strongest defining characteristic of samples in cluster 6 was high abundance of a few winter stonefly genera in the families Taeniopterygidae (Taeniopteryx, Strophopteryx, Taenionema) and Capniidae (Allocapnia). Prosimulium blackfly abundance was also fairly high in cluster 6 samples relative to other clusters, with samples in cluster 6 averaging 39 Prosimulium blackfly individuals per sample, the second highest of any cluster. So, cluster 6 can be characterized as representing streams – mostly of moderate size – sampled from early winter to early spring.

Samples in clusters 1-4 were from mostly moderate to larger, third to sixth order streams sampled at different times of the year. Broadly, samples in these clusters exhibited relatively high diversity and abundances of a few different mayfly genera as well as a few genera of Perlidae stoneflies. Since the fourth break in the cluster tree separated clusters 1-2 from clusters 3-4, discussion of these four clusters will focus on the fourth break.

Samples in clusters 1-2 were from mostly moderate size, second through fifth order streams while samples in clusters 3-4 were from mostly larger, fourth through sixth order streams. Samples in cluster 1 were collected mostly in April and May, while samples in cluster 2 were collected mostly November through May. Samples in cluster 3 were collected mostly August through October, but ranging from March through December, while samples in cluster 4 were almost all collected August through November. So, the break between clusters 1-2 and clusters 3-4 appears to be related to both drainage area and sampling season, with samples in clusters 1-2 mostly from moderate size streams sampled early winter through spring and samples in clusters 3-4 mostly from larger streams sampled mostly August through November. Notable differences in taxa between samples in clusters 1-2 and samples in clusters 3-4 are summarized above, but a broad pattern driving the split of these two groups of samples was the lower stonefly diversity in clusters 3-4, particularly in families other than Perlidae.

Discriminant Function Analysis

Discriminant function analysis (Fisher 1936; Hand 1981), another multivariate statistical technique, was used to further explore the results of the cluster analysis. This technique can be used to determine how much various parameters contribute to the classifications resulting from the cluster analysis.

A nonparametric linear discriminant function based on the four nearest-neighbors method using the 11 clusters from the cluster analysis as groups and 17 variables (Table 14) showed that stream size (measured by Strahler stream order and upstream drainage area in square miles) had the strongest coefficient values for the primary canonical function. Note that the primary canonical function had an eigenvalue nearly four times that of the secondary function, so most of the variability in the data was explained by the primary canonical function. Stream slope also showed a strong coefficient value for the primary canonical function, which is not surprising since it was highly correlated with stream size (Figure 5). Three percentage land use metrics (% forest, % developed, % agriculture) also showed fairly strong coefficient values for the primary canonical function, which can also be attributed in large part to correlations with stream size (i.e., even the most pristine larger basins have proportionally less forested land and more agricultural and developed land than smaller basins). Sampling season (measured either by Julian day or by month) showed fairly strong coefficients for the primary canonical function. Note that sampling month and Julian day are cyclical variables, so a linear model will not fully account for variability in these variables, although will provide some useful information. Other variables with fairly strong coefficients for the primary canonical function were the % calcareous geologies metrics (especially tertiary calcareous geologies) and total habitat score. The weakest primary function coefficients were for % wetland, latitude, and % abandoned mine lands. The primary function coefficients were also fairly weak for elevation and longitude. Looking at the secondary function, Julian day and month have the strongest coefficients by far, with latitude, slope, elevation, and % agriculture having moderately strong secondary function coefficients. Thus, the discriminant function analysis results suggest primacy of stream size and secondary importance of sampling season in determining the cluster groups.

Table 14. Canonical function coefficient values from a nonparametric linear discriminant function analysis using the 11 clusters from the cluster analysis as groups..

Variable	Coefficient value				
	Canonical function 1	Canonical function 2	Canonical function 3	Canonical function 4	Canonical function 5
Latitude	-0.08	-0.41	0.07	0.51	0.01
Longitude	0.19	0.06	0.35	-0.63	0.09
Julian Day	0.42	0.80	-0.01	0.18	0.11
Month	0.42	0.80	0.00	0.18	0.10
Drainage area	0.69	-0.03	-0.34	-0.03	-0.49
Strahler stream order	0.89	-0.14	-0.02	0.14	0.03
Slope	-0.64	0.30	0.05	-0.02	-0.42
Elevation	-0.17	0.30	-0.09	0.51	0.12
% primary calcareous geologies	0.29	-0.19	-0.07	-0.16	0.13
% secondary calcareous geologies	0.28	-0.05	-0.12	-0.15	0.19
% tertiary calcareous geologies	-0.38	0.15	-0.67	0.23	0.10
Total habitat score	-0.37	-0.08	0.02	-0.06	0.00
% forest	-0.48	0.16	0.12	0.39	-0.31
% developed	0.39	-0.08	-0.17	-0.41	0.04
% wetland	0.03	-0.03	-0.15	-0.35	0.40
% agriculture	0.58	-0.29	-0.07	0.10	0.00
% AML	0.08	0.10	0.01	-0.02	0.17
Eigenvalue	*2.19*	*0.57*	*0.46*	*0.21*	*0.20*
Cumulative proportion	*0.56*	*0.71*	*0.83*	*0.88*	*0.94*

Nonmetric Multidimensional Scaling

The same dataset that was used in the cluster analysis (i.e., the 192 most common taxa in 923 samples from 715 "condition 1" and "condition 2" sites) was run through NMDS analyses to look at the data from different perspectives. A variety of grouping and classification approaches – based on sampling season, stream size, and some other variables – were used in the NMDS analyses. The "badness-of-fit" – or "final stress" – criterion for the two-dimensional NMDS was 0.23.

Methods described by Van Sickle and Hughes (2000) – aided by MEANSIM, Version 6.0 software (Van Sickle 1998) – were used to quantify classification strengths of various grouping approaches (see Hawkins and Norris 2000). The classification strength of each grouping was quantified in two primary parameters: $\mathbf{W_{bar}}$, which measures within-group similarity; and $\mathbf{B_{bar}}$, which measures among-group similarity. Stronger grouping approaches minimize similarity among groups while maximizing similarity within groups, resulting in relatively low values of $\mathbf{B_{bar} / W_{bar}}$ and relatively high values of $\mathbf{W_{bar} - B_{bar}}$.

Since the groups from the cluster analysis can be thought of as an *a posteriori* classification scheme based solely on characteristics of the biological community, the goal of the a priori grouping approaches tested in the NMDS and classification strength analyses is to find the grouping approach or approaches – based on abiotic characteristics such as stream size and sampling season, for example – that result in classification strengths near that of the cluster analysis groupings. Because the preceding analyses show that stream size and sampling season most strongly influenced the cluster analysis groupings, the classification strength analyses focused on grouping approaches based on those factors.

Of the ten stream size grouping approaches examined (Table 15), the two approaches producing the strongest classifications were: a two-group approach based on drainage area (< 50 square miles and > 50 square miles) followed by a two-group approach based on Strahler stream order (1st to 4th order and 5th to 7th order). Another two-group approach based on drainage area (< 25 square miles and > 25 square miles) also produced a relatively strong classification, as did a three-group drainage area approach (< 25 square miles, 25 to 50 square miles, > 50 square miles). These findings reflect patterns seen in the first break in the cluster tree with 95% of the samples in clusters 8-11 coming from first, second, or third order sites mostly draining less than 25 to 50 square miles of land.

Of the five sampling season grouping approaches examined, the approach with the strongest classification was a four-group approach based on sample month (January to March, April to May, June to September, October to December). The sampling season grouping approach with each month as a group also had a fairly strong classification strength among the seasonal grouping approaches, but the aforementioned three-group seasonal approach resulted in a relatively high $\mathbf{W_{bar}-B_{bar}}$ value and a relatively low $\mathbf{B_{bar}/W_{bar}}$ value while the 12-group monthly approach had a relatively high $\mathbf{W_{bar}-B_{bar}}$ value but fairly high $\mathbf{B_{bar}/W_{bar}}$ value. Two other grouping approaches based on sampling month produced fairly strong classifications among the seasonal grouping approaches: another four-season approach (March to May, June to August, September to November, December to February) and a three-season approach (March to May, June to September, October to February). Nearly all the grouping approaches based on sampling seasons were weaker than approaches based on stream size, further confirming the primary influence of stream size

and secondary influence of sampling season on patterns observed in the most common taxa in samples from “condition 1” and “condition 2” sites.

Six grouping approaches were examined based on combinations of the strongest stream size and a few of the strongest sampling season grouping approaches. Of these six grouping approaches, the strongest classification strengths were produced by two four-group approaches based of two stream size groups (< 50 square miles and > 50 square miles for drainage area; 1st to 4th order and 5th to 7th order for Strahler stream order) and two seasonal groups (October to May and June to September). Grouping approaches based on both stream size and sampling season tended to be stronger than the corresponding approaches based only on sampling season, but weaker than the corresponding approaches based only on stream size. Interestingly, the seasonal grouping approaches that produced the strongest classifications did not result in the strongest classifications when combined with stream size components, rather the two-season (October to May and June to September) grouping approaches produced stronger classifications when combined with stream size factors. Some of this phenomenon may be attributable to limited sampling of larger streams during winter months (e.g., there were only 7 samples from 5th order or larger streams in the January to March season).

Overall, the two grouping approaches that produced the strongest classifications – aside from the cluster tree groupings – were based on stream size. These two grouping approaches – along with patterns observed in the cluster analysis – suggest that the most significant taxonomic patterns in samples from relatively undisturbed sites relate to stream size, with strong differences in taxa abundances and occurrences between first, second, third, and fourth order streams draining less than 25 or 50 square miles of land, and fifth, sixth, and seventh order streams draining more than 25 or 50 square miles of land.

The patterns and performance of each grouping approach can be visualized in NMDS ordination plots with symbols coded according to the various groups. Only two NMDS ordinations are reproduced here: the cluster analysis groupings (Figure 19); and the four-group approach based on two drainage area groups (< 50 square miles and > 50 square miles) and two seasonal groups (October to May and June to September) (Figure 20). The strongest patterns in the NMDS ordinations are along the Dimension 1 axes with secondary patterns plotted along the Dimension 2 axes. Notice that the NMDS ordination of the cluster groups shows similar patterns as the cluster tree, with fairly distinct groupings of clusters 1-4, 8-10, and 11 along the Dimension 1 axis and further separation of clusters 1-2 from clusters 3-4 as well as clusters 5-7 from clusters 1-2 and 8-10 along the Dimension 2 axis. Note that the polarity of the Dimension 1 axes in the two NMDS ordinations (Figure 19, Figure 20) are reversed from each other (i.e., spring samples from small streams appear to the left in Figure 19, but to the right in Figure 20). This is an artifact of the way the plots are produced. The designation of one end of each NMDS axis as positive and one as negative is arbitrary. What is important is the relative position of samples in the plots to one another.

It is not surprising that none of the abiotic grouping approaches produced classifications as strong as that of the cluster analysis groupings since there was substantial overlap among the eleven clusters for many abiotic parameters (e.g., drainage area, stream order, slope, elevation, sampling season, basin, latitude, longitude). Still, the NMDS and classification strength analyses further support that stream size is a primary variable driving taxonomic patterning in relatively undisturbed streams, with secondary influence of sampling seasons.

Table 15. Classification strengths of various grouping approaches of the 192 most common taxa in 923 samples from 715 reference sites. After Table 1 of Van Sickle and Hughes (2000).

Grouping Basis	Groups	# of groups	W_{bar} (%)	B_{bar} (%)	W_{bar} - B_{bar} (%)	B_{bar} / W_{bar} (%)
Cluster analysis						
Clusters	(1) (2) (3) (4) (5) (6) (7) (8) (9) (10) (11)	11	51.5	36.9	**14.6**	**71.7**
Stream size						
Strahler Stream order	(1) (2) (3) (4) (5) (6-7)	6	42.4	35.4	7.0	83.4
	(1) (2-3) (4) (5-7)	4	41.7	33.8	8.1	78.6
	(1-3) (4) (5-7)	3	41.0	32.4	8.6	79.0
	(1-3) (4-7)	2	40.0	32.8	7.2	81.9
	(1-4) (5-7)	2	39.8	29.2	**10.5**	**73.5**
Drainage area (square miles)	(0-3) (3-10) (10-25) (25-50) (50-100) (100-500) (500-1,000)	7	42.8	34.4	8.4	80.4
	(0-10) (10-25) (25-100) (100-1,000)	4	41.8	32.8	9.0	78.4
	(0-25) (25-50) (50-1,000)	3	40.4	30.0	10.4	74.3
	(0-25) (25-1,000)	2	40.1	29.6	10.5	73.7
	(0-50) (50-1,000)	2	39.8	28.9	**10.9**	**72.6**
Sampling season						
Months	(Jan) (Feb) (Mar) (Apr) (May) (Jun) (Jul) (Aug) (Sep) (Oct) (Nov) (Dec)	12	45.1	38.1	7.0	84.5
	(Jan-Mar) (Apr-May) (Jun-Sep) (Oct-Dec)	4	44.1	36.7	**7.4**	**83.2**
	(Mar-May) (Jun-Aug) (Sep-Nov) (Dec-Feb)	4	42.8	36.0	6.8	84.1
	(Mar-May) (Jun-Sep) (Oct-Feb)	3	42.7	35.8	6.9	83.8
	(Oct-May) (Jun-Sep)	2	40.7	35.0	5.7	86.0
Stream size x Sampling season						
Strahler Stream Order x Months	(1-4) (5-7) x (Jan-Mar) (Apr-May) (Jun-Sep) (Oct-Dec)	8	39.9	31.8	8.1	79.8
	(1-4) (5-7) x (Mar-May) (Jun-Sep) (Oct-Feb)	6	39.2	30.6	8.6	78.0
	(1-4) (5-7) x (Oct-May) (Jun-Sep)	4	37.8	28.0	**9.8**	**74.0**
Drainage area (square miles) x Months	(0-50) (50-1,000) x (Jan-Mar) (Apr-May) (Jun-Sep) (Oct-Dec)	8	40.4	31.8	8.5	78.9
	(0-50) (50-1,000) x (Mar-May) (Jun-Sep) (Oct-Feb)	6	39.7	30.7	9.1	77.2
	(0-50) (50-1,000) x (Oct-May) (Jun-Sep)	4	38.3	28.3	**9.9**	**74.0**

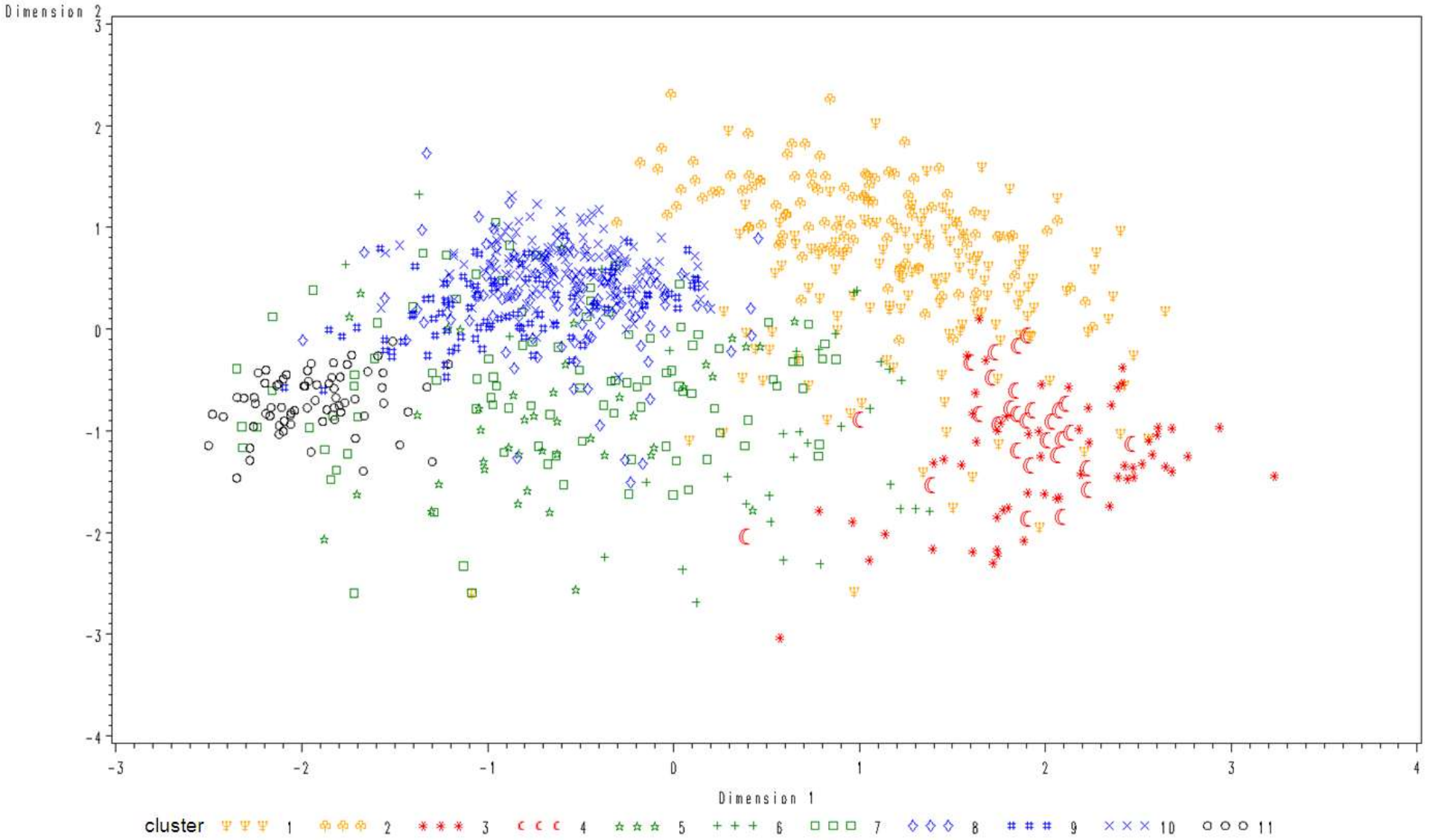


Figure 19. NMDS ordination plot (first two dimensions) of the 192 most common taxa in 923 samples from 715 reference sites, coded by the 11 clusters from the cluster analysis.

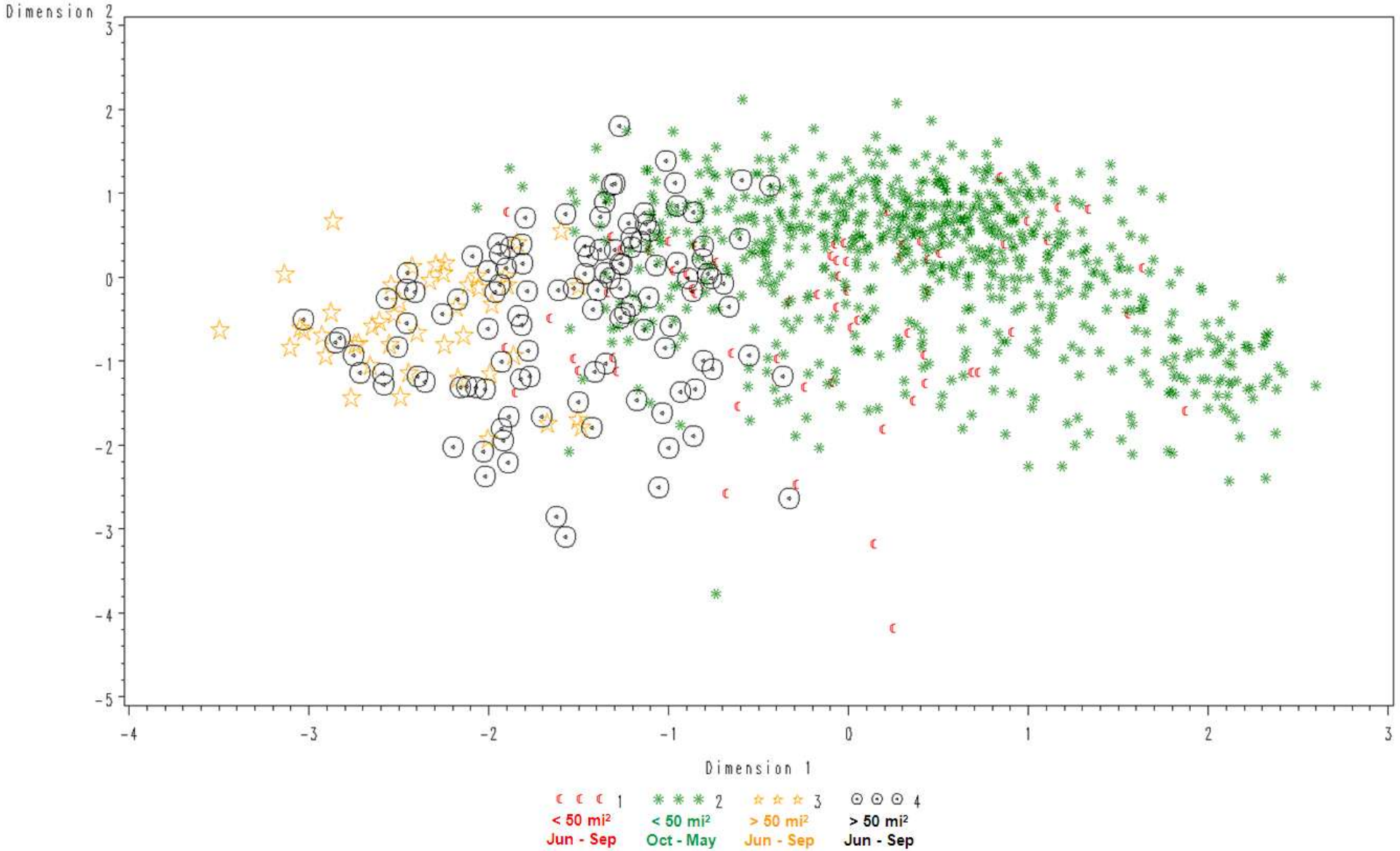


Figure 20. NMDS ordination plot (first two dimensions) of the 192 most common taxa in 923 samples from 715 reference sites, coded by drainage area range (< 50 square miles; > 50 square miles) and sampling seasons (June to September; October to May).

METRICS ANALYSIS AND INDEX DEVELOPMENT

A biological metric quantifies measurable characteristics of the biota that changes in predictable ways with increased anthropogenic stress (Barbour et al. 1995). Metrics measure meaningful indicator attributes in assessing the biological condition of sample sites (Barbour et al. 1999). Vast arrays of metrics have been tested in developing various indices of biotic integrity for a variety of aquatic assemblages, including benthic macroinvertebrates (Barbour et al. 1995). The utility of each metric is based on a hypothesis about the predictable relationship between the biological response measured by that metric and ecosystem stress caused by human impacts (Barbour et al. 1995; Yoder and Rankin 1995).

Multimetric Indices

Most water resource agencies in the United States use a multimetric approach to developing indices of biological integrity (Barbour et al. 1999). This approach utilizes a suite of metrics that measure diverse biological attributes and respond to different stressors. A major benefit of the multimetric approach is the ability to incorporate information from a number of metrics that, when integrated into a single measure, or index, can provide a meaningful indicator of overall biological condition (Barbour et al. 1995). Such an index helps to increase sensitivity to a broad range of ecosystem stressors and to minimize any weaknesses or limitations that each underlying metric may have if used individually. For example, some metrics are sensitive across a broad range of biological conditions and other metrics are only sensitive in part of the range. Metrics that exhibit detectable responses to changing disturbance conditions are important for indicating comparability to – or departure from – the established reference biological condition. Overlap in the ranges of sensitivity of individual metrics helps strengthen conclusions regarding biological condition reached using an integrative, multimetric index approach (Barbour et al. 1995).

Pollution Tolerance

PADEP assigns numeric pollution tolerance value (PTV) to most benthic macroinvertebrate taxa encountered in Pennsylvania. These PTVs are integer values that range from zero to ten, with values closer to zero representing relative sensitivity to pollution and values closer to ten representing relative tolerance of pollution. The PTVs are based on information from a number of sources, such as U.S. EPA (1990) and Barbour et al. (1999). Experience and many studies indicate that different organisms respond differently to different types of pollution (e.g., Pond 2010; Carlisle et al. 2007). Most of the PTVs used by PADEP to date reflect organismal responses to pollution related to organic enrichment and sedimentation, while these PTVs are not necessarily reflective of organismal responses to other types of pollution, notably low pH conditions related to stream acidification. For example, Leuctra stoneflies are assigned a PTV of 0, reflecting the extreme sensitivity of this stonefly genus to organic pollution and sedimentation. However, Leuctra stoneflies are very tolerant of low pH conditions and often thrive in streams that experience low pH conditions (Madarish and Kimmel 2000; Kimmel 1999; Rosemond et al. 1992; Simpson et al. 1985). Similarly, Baetis mayflies are assigned a PTV of 6, indicating relative tolerance of organic pollution and sedimentation, but this mayfly genus – as with most mayfly taxa (Madarish and Kimmel 2000; Kimmel 1999; Rosemond et al. 1992; Simpson et al. 1985) – are quite sensitive to low pH conditions.

PADEP also assigned a second set of numeric values to most taxa related to pollution sensitivity and tolerance. This second set of values – named as Biological Condition Gradient (BCG) attributes – was assigned at a series of Tiered Aquatic Life Use (TALU) workshops, which are further described below, by Gerritsen and Jessup (2007), and by Davies and Jackson (2006). Davies and Jackson (2006) provide more detail on the BCG attribute values, but briefly, these BCG attribute values are integers that range from one to six with the six attributes described as: (I) historically documented, sensitive, long-lived, or regionally endemic taxa; (II) sensitive-rare taxa; (III) sensitive-ubiquitous taxa; (IV) taxa of intermediate tolerance; (V) tolerant taxa; and (VI) non-native or intentionally introduced taxa. Like the PTVs, the BCG attribute values are mostly geared towards organismal responses to pollution related to organic enrichment and sedimentation and do not always reflect responses to other types of pollution, notably acidification.

Since pH was not recorded for 887 of the 2,482 sites, and since stream acidification can be a seasonal phenomenon that one-time pH observations may not pick up, the dataset was screened for samples exhibiting impacts from acid deposition using the biological patterns seen in samples in cluster 11 of the cluster analysis. Any sample with less than 5% mayfly individuals combined with over 25% Amphinemura and/or Leuctra individuals was flagged as likely impacted by acid deposition. For metrics analysis and index development, these samples were grouped with samples that had pH values recorded below 5.5 into an “acid impacted” category for each stream size class as defined above. Although classifying samples based on biology is not typically done in developing indices of biological integrity, the paucity of available pH data and ephemeral nature of acid deposition impacts makes it difficult to address these impacts through abiotic parameters. Plus, there is strong empirical evidence and support in the literature characterizing acid deposition impacted benthic macroinvertebrate communities. If this was not done, a substantial number of acid impacted samples would have been grouped in with ‘condition 1” and “condition 2” samples, unduly skewing the reference conditions.

Since some metrics are based on PTVs and/or BCG attributes, the types of pollution they are calibrated to carry implications for interpreting metric and multimetric indices. These implications are discussed in more detail below.

Candidate Metrics

Ideally, evaluation of candidate metrics should result in selection of metrics that: (1) are based in well-understood ecological principles relevant to the biological community in the type of water body being studied as well as to sampling methods and assessment objectives; (2) respond to anthropogenic stress in a predictable manner; (3) have responses to stressors that can be distinguished from natural variation and that can discriminate along a gradient of anthropogenic stress; (4) are environmentally benign to measure; and (5) are cost-effective to sample (Barbour et al. 1995). Barbour et al. (1999) and Flotemersch et al. (2006) offer additional relevant considerations for selecting metrics. The most useful indices of biological integrity incorporate metrics based on sound ecological principles and representing diverse aspects of structure, composition, individual health, and/or processes of the biological community. Such metrics quantify expectations defined by the reference condition and can serve as the foundation of a sound, integrated assessment of biological condition (Barbour et al. 1995).

A number of major classes of attributes have been generally defined for metrics applied to benthic macroinvertebrate communities: taxonomic richness; community composition; pollution tolerance; trophic guild; behavior or motility habit; and life cycle (Barbour et al. 1999). Candidate metrics considered in this analysis generally fit into one of these major categories, although some metrics incorporate aspects of two or more of these major classes (Table 16). No measures of individual condition were considered because PADEP does not routinely assess nor record individual condition of benthic macroinvertebrates.

Table 16. Candidate metrics evaluated in this project.

Taxa	Taxa Richness	Proportional Taxa Richness	% Individuals	Other	Notes	Expected Response to Increasing Anthropogenic Stress
Total Taxa	X					Decrease
Mayfly Taxa ***	X	X	X			Decrease
Stonefly Taxa	X	X	X			Decrease
Caddisfly Taxa ***	X	X	X			Decrease
Mayfly (E) + Stonefly (P) + Caddisfly (T) Taxa***	X	X	X			Decrease
BCG Attribute I Taxa	X	X	X			Decrease
BCG Attribute II Taxa	X	X	X			Decrease
BCG Attribute III Taxa	X	X	X			Decrease
BCG Attribute I + II + III Taxa	X	X	X			Decrease
BCG Attribute IV Taxa	X	X	X			Increase
BCG Attribute V Taxa	X	X	X			Increase
BCG Attribute IV + V + VI Taxa	X	X	X			Increase
(BCG Attribute I + II + III Taxa) / (BCG Attribute IV + VI + VI Taxa)	X	X	X			Decrease
PTV 0 – 5 Taxa	X	X	X			Decrease
PTV 0 – 4 Taxa	X	X	X			Decrease
PTV 0 – 3 Taxa	X	X	X			Decrease
PTV 0 – 2 Taxa	X	X	X			Decrease
PTV 5 – 10 Taxa	X	X	X			Increase
PTV 6 – 10 Taxa	X	X	X			Increase
PTV 7 – 10 Taxa	X	X	X			Increase
PTV 8 – 10 Taxa	X	X	X			Increase
Hilsenhoff Biotic Index				X	Number of individuals weighted by PTV score	Increase
BCG Index				X	Number of individuals weighted by BCG Attribute	Increase
Beck's Index				X	Taxa richness weighted by PTV – multiple versions tested	Decrease
Predator Taxa	X	X	X			Decrease
Shredder Taxa	X	X	X			Decrease
Filter-Collector Taxa	X	X	X			Increase
Collector-Gatherer Taxa	X	X	X			Increase
Scraper Taxa	X	X	X			Increase
Dominant Taxa			X			Increase
Shannon Diversity				X	Distribution of individuals among taxa	Decrease
Non-Insect Taxa			X			Increase
Oligochaeta Taxa			X			Increase
Diptera Taxa ***	X		X			Variable
Chironomidae Taxa			X			Increase
Hydropsychidae Taxa ***	X	X	X		Used various combinations of genera	Increase

*** these metrics were computed using all taxa and using only certain sensitive and/or tolerant taxa

Discrimination Efficiency

As a first cut, the ability of each candidate metric to discriminate along a gradient of anthropogenic impacts was evaluated visually by looking at boxplots of values of each candidate metric by the condition categories defined above.

For metrics that exhibited ability to distinguish conditions along this gradient of impact severity, discrimination efficiencies were calculated in order to quantify the ability of each metric to distinguish least impacted from most impacted condition. For metrics expected to decrease in value with increasing anthropogenic stress, or negative-response metrics, the following equation was used to calculate the discrimination efficiency:

$$D.E.\ (\%) = \mathbf{n}_{\mathbf{condition6<\%condition1}} / \mathbf{n}_{\mathbf{condition6total}} * 100$$

where

- **D.E.** = the discrimination efficiency
- $\mathbf{n}_{\mathbf{condition6<\%condition1}}$ = the number of "condition 6" samples with metric values less than the 25th percentile value of all "condition 1" samples, and
- $\mathbf{n}_{\mathbf{condtion6total}}$ = the total number of "condition 6" samples.

For metrics expected to increase in value with increasing stress, or positive-response metrics, the following equation was used to calculate the discrimination efficiency:

$$D.E.\ (\%) = \mathbf{n}_{\mathbf{condition6>\%condition1}} / \mathbf{n}_{\mathbf{condition6total}} * 100$$

where

- **D.E.** = the discrimination efficiency
- $\mathbf{n}_{\mathbf{condition6>\%condition1}}$ = the number of "condition 6" samples with metric values greater than the 75th percentile value of all "condition 1" samples, and
- $\mathbf{n}_{\mathbf{condition6total}}$ = the total number of "condition 6" samples.

Metrics with minimal or no overlap between the distribution of scores for "condition 1" and "condition 6" samples (i.e., high discrimination efficiencies) can be considered strong, predictable discriminators between reference and stressed conditions. Such metrics provide the most confidence for assessing the biological condition of unknown sites (Barbour et al. 1999).

Discrimination efficiencies were evaluated within each stream size class as defined above. Metrics with high discrimination efficiencies were selected for further evaluation. With such a large number of metrics evaluated, discrimination efficiencies are presented below only for the six metrics selected for inclusion in the final IBI (Table 17). Discrimination efficiency evaluations for other candidate metrics are available upon request.

Table 17. Discrimination efficiencies of selected core metrics by drainage area range.

Metric	Expected Response to Increasing Anthropogenic Stress	Drainage area range (square miles)						
		0 to 3	3 to 10	10 to 25	25 to 50	50 to 100	100 to 500	500 to 1,000
		Discrimination Efficiency						
Total Taxa Richness	Decrease	87%	89%	89%	94%	94%	76%	56%
EPT Taxa Richness (PTV 0-4 only)	Decrease	97%	96%	97%	98%	100%	92%	81%
Beck's Index (version 3)	Decrease	98%	99%	98%	100%	100%	98%	100%
Hilsenhoff Biotic Index	Increase	90%	90%	93%	96%	94%	84%	88%
Shannon Diversity	Decrease	90%	94%	94%	94%	86%	70%	56%
% Sensitive Individuals (PTV 0-3 only)	Decrease	89%	90%	91%	92%	91%	82%	88%

Discrimination efficiencies were excellent across all stream sizes for most of the six core metrics. Discrimination efficiencies dropped off a bit for the Total Taxa Richness metric and the Shannon Diversity metric in streams draining more than 100 square miles of land, and especially in streams draining more than 500 square miles of land. These two metrics are fairly strongly correlated, and the reduced discrimination efficiencies for both metrics in larger streams can be explained by a few factors. Firstly, there are relatively few larger streams compared with the number of smaller streams. This means that we are comparing fewer samples from fewer sites as streams get larger and larger. For example, in the 7.6 condition category, there are 16 samples from 8 sites on 5 streams. In the 7.1 condition category there are 19 samples from 4 sites on 3 streams. So, we are comparing small numbers of samples and sites in these larger streams. A second factor reducing the discrimination efficiencies of these two diversity metrics in larger streams has to do with the nature of the human activities in the basins of the larger streams classified as condition category 7.6 in this dataset. The six samples with the highest taxonomic diversity (measured by either Total Taxa Richness or Shannon Diversity) were collected from two locations in French Creek in the northwest part of Pennsylvania. Agriculture occupies over 33% of the land use in this basin, but the in-stream impacts of human activities in this basin do not manifest as reduced overall taxonomic diversity. Rather, we see these impacts in the relative preponderance of more tolerant taxa (e.g., Simulium, Stenacron, Anthopotamus, Oligochaeta, Gammarus, gastropods, Sphaeriidae) compared with the taxa we see more in the 7.1 samples (e.g., Acroneuria, Leucrocuta, Serratella, Epeorus, Corydalus, Isonychia). A third factor likely contributing to the reduced discrimination efficiency of these two metrics is seasonality. These discrimination efficiencies were calculated with samples collected throughout the year. However, the 7.6 samples with the highest taxonomic diversity were collected in April, May, or November while the 7.1 samples with the lowest taxonomic diversity were collected in September or October, when we expect naturally lower seasonal diversity.

It can be argued that – due to their reduced discrimination efficiencies in larger streams – the Total Taxa Richness metric and the Shannon Diversity metric should be dropped from or replaced in the IBI for larger streams. However – even with these two metrics left in – the large-stream IBI demonstrated very good discrimination efficiency in streams of all sizes (as shown below). Furthermore, there is value to utilizing the same suite of metrics across the

board in terms of programmatic consistency and in terms of communicating assessment methods to colleagues and the public. With that in mind, IBI scores for larger streams should be interpreted mindful of these taxonomic diversity metric considerations.

Metric Correlations

In order to help select strongly discriminating metrics while reducing the number of metrics relating redundant information, metric correlations were analyzed for all metrics with high discrimination efficiencies. Due to the large number of metrics analyzed, correlations are presented here only for the six metrics selected for inclusion in the final IBI (Table 18). Correlation analyses for other candidate metrics are available upon request.

Table 18. Pearson correlation (r) values for selected core metrics.

Metric	Total Taxa Richness	EPT Taxa Richness (PTV 0-4 only)	Beck's Index (version 3)	Hilsenhoff Biotic Index	Shannon Diversity	% Sensitive Individuals (PTV 0-3 only)
Total Taxa Richness	1	----	---	---	---	---
EPT Taxa Richness (PTV 0-4 only)	0.87	1	---	---	---	---
Beck's Index (version 3)	0.75	0.91	1	---	---	---
Hilsenhoff Biotic Index	-0.48	-0.69	-0.73	1	---	---
Shannon Diversity	0.86	0.77	0.66	-0.49	1	---
% Sensitive Individuals (PTV 0-3 only)	0.43	0.67	0.71	-0.93	0.41	1

The correlation between the EPT Taxa Richness (counting only taxa with pollution tolerance values, or PTVs, 0 – 4) metric and the Beck's Index, version 3 metric was fairly high (r = 0.91), as was the correlation between the Hilsenhoff Biotic Index metric and the % Sensitive Individuals (PTV 0 – 3) metric (r = -0.93). However, scatterplots of the relationship between these pairs of metrics (Figure 21) revealed enough variation that all were retained for inclusion in the final IBI.

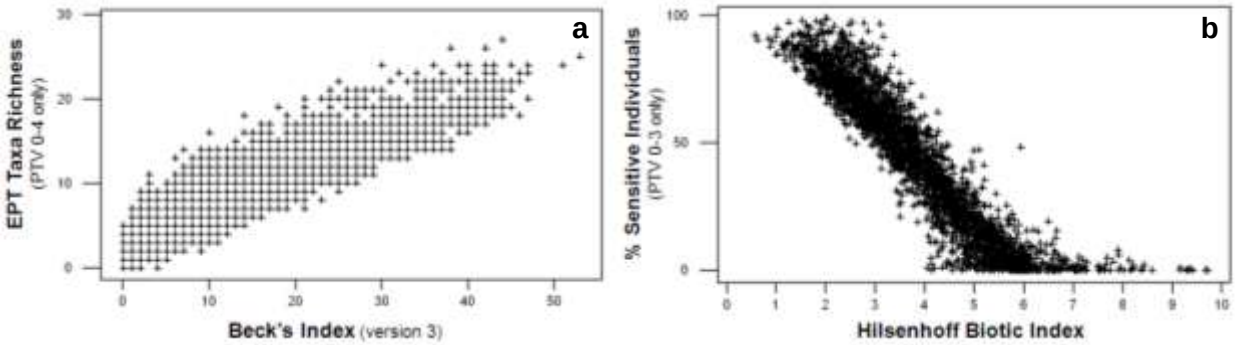


Figure 21. Scatterplots of (a) EPT Taxa Richness and Beck's Index metric scores and (b) Percent Sensitive Individuals and Hilsenhoff Biotic Index metric scores.

Core Metrics

A number of different metric combinations were evaluated during index development. Based on discrimination efficiencies, correlation matrix analyses, and other index

performance characteristics discussed below, the following six metrics were selected for inclusion as core metrics in the multimetric index (Appendix C shows examples of the six core metric and index calculations for a sample and Appendix D contains the pollution tolerance values for all taxa in this dataset).

1. Total Taxa Richness

This taxonomic richness metric is a count of the total number of taxa in a sub-sample. Generally, this metric is expected to decrease with increasing anthropogenic stress to a stream ecosystem, reflecting loss of taxa and increasing dominance of a few pollution-tolerant taxa. Other benefits of including this metric include its common use in many biological monitoring and assessment programs in other parts of the world as well as its ease of explanation and calculation.

2. Ephemeroptera + Plecoptera + Trichoptera Taxa Richness (PTV 0-4 only)

This taxonomic richness metric is a count of the number of taxa belonging to the orders Ephemeroptera, Plecoptera, and Trichoptera (EPT) in a sub-sample – common names for these orders are mayflies, stoneflies, and caddisflies, respectively. The aquatic life stages of these three insect orders are generally considered sensitive to, or intolerant of, many types of pollution (Lenat and Penrose 1996), although sensitivity to different types of pollution varies among taxa in these insect orders. The version of this metric used here only counts EPT taxa with PTVs of 0 to 4, excluding a few of the most tolerant mayfly and caddisfly taxa. This metric is expected to decrease in value with increasing anthropogenic stress to a stream ecosystem, reflecting the loss of taxa from these largely pollution-sensitive orders. This metric has a history of use across the world and is relatively easy to use, explain, and calculate (Lenat and Penrose 1996).

3. Beck's Index (version 3)

This taxonomic richness and tolerance metric is a weighted count of taxa with PTVs of 0, 1, or 2. The name and conceptual basis of this metric are derived from the water quality work of William H. Beck in Florida (Beck 1955). This metric is expected to decrease in value with increasing anthropogenic stress to a stream ecosystem, reflecting the loss of pollution-sensitive taxa. It should be noted that the version of the Beck's Index metric used for this project, although similar in name and concept, differs slightly in its calculation from the Beck's Index used in PADEP's multihabitat protocol for assessing biological condition of low gradient, pool-glide type streams (see Appendix C for calculation details).

4. Shannon Diversity

This community composition metric measures taxonomic richness and evenness of individuals across taxa in a sub-sample. This metric is expected to decrease in value with increasing anthropogenic stress to a stream ecosystem, reflecting loss of pollution-sensitive taxa and increasing dominance of a few pollution-tolerant taxa. The name and conceptual basis for this metric are derived from the information theory work of Claude Elwood Shannon (Shannon 1968).

5. Hilsenhoff Biotic Index

This community composition and tolerance metric is calculated as an average of the number of individuals in a sub-sample, weighted by PTVs. Developed by William Hilsenhoff, the Hilsenhoff Biotic Index (Hilsenhoff 1977, 1987, 1988; Klemm et al. 1990) generally increases with increasing ecosystem stress, reflecting increasing dominance of pollution-tolerant organisms.

6. Percent Sensitive Individuals (PTV 0-3 only)

This community composition and tolerance metric is the percentage of individuals with PTVs of 0 to 3 in a sub-sample and is expected to decrease in value with increasing anthropogenic stress to a stream ecosystem, reflecting loss of pollution-sensitive organisms.

These six metrics all exhibited a strong ability to distinguish between reference and stressed conditions. In addition, these six metrics measure different aspects of the biological communities represented by the sub-samples. When used together in a multimetric index, these metrics provide a solid foundation for assessing the biological condition of benthic macroinvertebrate assemblages in Pennsylvania's wadeable, freestone, riffle-run stream ecosystems.

A number of different metric combinations were evaluated during index development and that this combination of metrics provided among the strongest performance characteristics of any metric combination tested. The selected six metrics do not include a metric that directly utilizes the functional feeding group assignment of each taxon (e.g., scraper, predator, shredder). A functional feeding metric was not included in the multimetric index for a number of reasons, primarily because of the difficulty predicting how functional feeding metrics respond to different anthropogenic stressors and because natural changes are expected in the distribution of organisms among functional feeding groups with increasing drainage area and associated changes in a stream's trophic dynamics (Vannote et al. 1980). These factors limit the range of applicability of functional feeding metrics to certain stream sizes; further, difficulties with proper assignment of taxa to functional feeding groups contribute to the unreliability of these metrics.

Core Metrics, Stream Size, Sampling Season

Since the above analyses show that benthic macroinvertebrate communities in relatively undisturbed streams naturally vary with stream size and sampling season, PADEP thought it prudent to consider how these natural variations manifest in the selected core metrics.

If we look just at "condition 1" samples to minimize the influence of anthropogenic impacts, some of the selected core metrics exhibit distinct patterns with stream size (Figure 22). The Beck's Index metric displays the strongest variability with stream size, with higher values observed in samples from small streams than from larger streams. The Hilsenhoff Biotic Index and the Percent Sensitive Individuals metrics also exhibit fairly strong correlations with stream size, with lower Hilsenhoff Biotic Index values and higher Percent Sensitive Individuals values in samples from smaller streams than from larger streams. The Total Taxa Richness and the EPT Richness metrics display weaker, although still noticeable variability with stream size, with lower values observed in samples from larger streams than from smaller streams. This pattern is weaker for the Total Taxa Richness metric than the EPT Richness metric. The Shannon Diversity metric does not vary much with stream size.

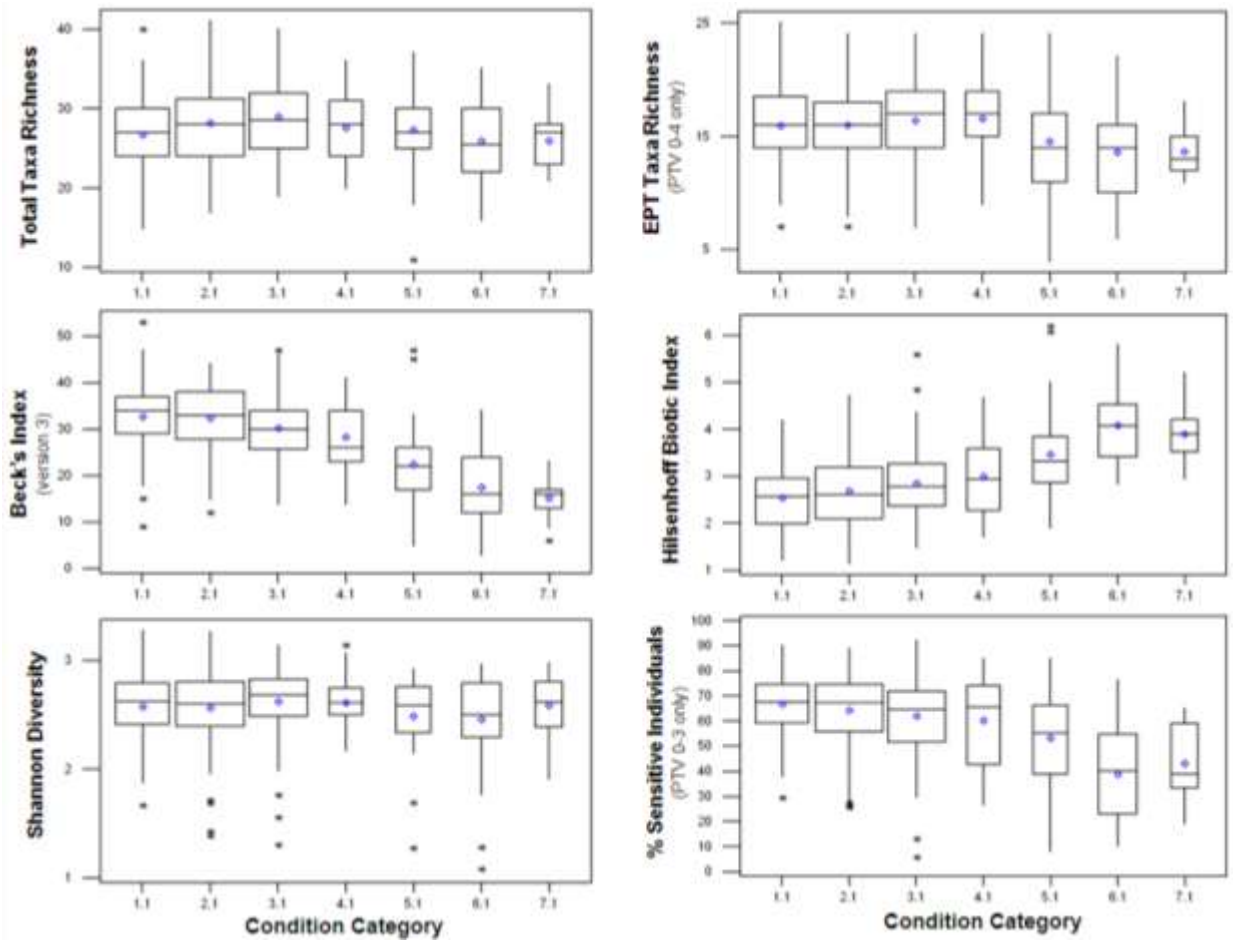


Figure 22. Boxplots of core metric distributions for "condition 1" samples by size range. The first digit in each condition category represents square-mile drainage area ranges (1 = 0 to 3 mi^2; 2 = 3 to 10 mi^2; 3 = 10 to 25 mi^2; 4 = 25 to 50 mi^2; 5 = 50 to 100 mi^2; 6 = 100 to 500 mi^2; 7 = 500 to 1,000 mi^2). The second condition category digit indicates "condition 1" designation.

The correlation of these metric scores with stream size can be explained by the patterns seen in the cluster analysis. Taxa with very low PTVs (i.e., 0 or 1) – such as: Ameletus, Paraleptophlebia, Epeorus, and Cinygmula, mayflies; Pteronarcys, Tallaperla, Leuctra, Haploperla, Alloperla, and Sweltsa stoneflies; Wormaldia, Dolophilodes, Parapsyche, Diplectrona, and Rhyacophila caddisflies – are much less commonly encountered in larger streams than smaller streams. Rather, we more commonly encounter taxa with higher PTVs (i.e., >2) – such as: Isonychia, Acentrella, Plauditus, Maccaffertium, Stenonema, and Caenis mayflies; Cheumatopsyche, Ceratopsyche, Hydropsyche, Macrostemmum, and Chimarra caddisflies. That is not to say we don't encounter low PTV taxa – such as: Heterocloeon, Leucrocuta, and Serratella mayflies; and Acroneuria and Paragnetina stoneflies – in larger streams, but such taxa typically compose a much smaller proportion of the taxa and individuals in larger streams than smaller streams. These patterns have the greatest impact on metrics that are weighted by PTVs: the Beck's Index metric, which is a taxa richness based metric weighted by PTVs; and the Hilsenhoff Biotic Index metric, which is an abundance based metric weighted by PTVs. The patterns of occurrence and abundance of low PTV taxa also impact – although to a lesser extent than PTV-weighted metrics – the metrics that only count lower PTV taxa: the EPT Taxa Richness metric, which

only counts EPT taxa with PTVs less than 5; and the Percent Sensitive Individuals metric, which only counts individuals from taxa with PTVs less than 4. The Total Taxa Richness metric and Shannon Diversity metric do not show much variation with stream size – particularly the Shannon Diversity metric – in part because these metrics do not incorporate PTVs into their calculation. The Total Taxa Richness metric shows a slight drop with increasing stream size mainly because – broadly speaking – we often see reduced diversity in the stonefly (Plecoptera) and true fly (Diptera) orders in larger streams, but this is somewhat tempered because we often see increased diversity of mayflies (Ephemeroptera) in larger streams. Keep in mind, these patterns are described using the target taxonomic levels utilized by PADEP (e.g., family level for Chironomidae).

If we look just at "condition 1" samples to minimize the influence of anthropogenic effects, some of the selected core metrics exhibit distinct patterns with sampling season (Figure 23). Although "condition 1" samples from June and July are rare, the three metrics based on taxa richness (Total Taxa Richness, EPT Taxa Richness, Beck's Index) exhibit substantial drops in scores during the summer and early autumn, from June through October. The Hilsenhoff Biotic Index metric and the Percent Sensitive Individuals metric also show distinct patterns in scores during the summer and early autumn. Since these two metrics respond oppositely to increasing anthropogenic stress, the patterns sort of mirror one another, but exhibit similar seasonality. Beginning in June – possibly even late May – the Percent Sensitive Individuals scores drop and the Hilsenhoff Biotic Index scores rise. In September and October, the Percent Sensitive Individuals scores begin to rise again and the Hilsenhoff Biotic Index scores begin to drop again, although these metrics do not appear to return to close to their respective maximum and minimum potentials until November. The Shannon Diversity metric exhibits only a slight drop in scores during the summer and early autumn relative to the other five metrics. Scores for this metric also return to full maximum potential in November.

The seasonal patterns in metric scores can be explained by the phenological life cycles of many benthic macroinvertebrate taxa. For example, mayflies – as a taxonomic order – are named as such because many species in this order emerge from streams as subimagos in the month of May. Of course, different mayfly species exhibit different life cycle characteristics and timing – some mayflies emerge from streams in late March, some in September. The preceding characterization of many mayflies emerging from streams in the month of May was just a simple, low-hanging example meant to illustrate that many benthic macroinvertebrate life cycles follow predictable seasonal cycles. During the summer and early autumn months, many taxa are present in stream benthos in egg stages or very early – and small – instars. During these times of the year, we often observe reduced benthic diversity because we cannot easily identify these organisms in such miniscule life stages. As autumn ends and winter arrives, many of these organisms become large enough to accurately and precisely identify. Diversity can also be observed to increase in the winter months as winter stonefly taxa mature and other taxa continue to grow. This is a very broad characterization of phenological phenomena observed in benthic macroinvertebrates. Each species of organism exhibits different, nuanced life cycle patterns – a treatment of which is beyond the scope of this report – but these broad observations help explain why we see metrics behave as they do with changing seasons of the year.

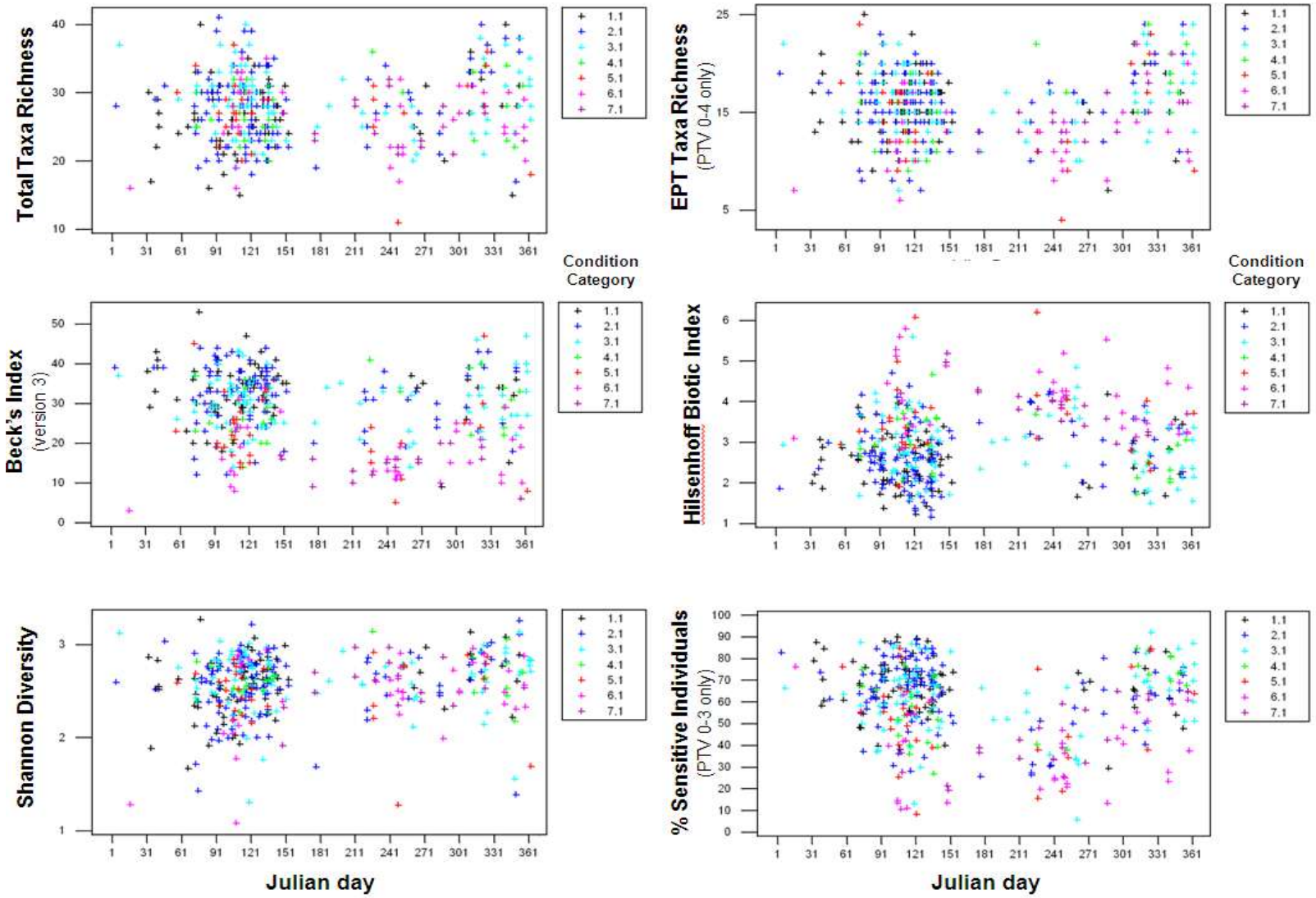


Figure 23. Scatterplots of core metrics by Julian day of sample collection and color-coded by condition category. The first digit in each condition category represents drainage area ranges (1 = 0 to 3 square miles; 2 = 3 to 10 square miles; 3 = 10 to 25 square miles; 4 = 25 to 50 square miles; 5 = 50 to 100 square miles; 6 = 100 to 500 square miles; 7 = 500 to 1,000 square miles). The second condition category digit indicates "condition 1" designation.

Index Development

An index is simply a means to integrate information from various metrics of biological integrity (Barbour et al. 1999). In order to compare and combine sundry measures (e.g., percentage of individuals, counts of taxa, unitless numbers) of biological condition in a meaningful manner, it is necessary to standardize metrics with some mathematical transformation that results in a logical progression of values (Barbour et al. 1995). Barbour et al. (1999) recommend using a composite of sites representing a gradient of biological conditions (e.g., natural to severely degraded) in the metric standardization and index development process to calibrate the index to a range of biological conditions.

Each selected core metric was evaluated at a selected percentile of the distribution of all samples by the size groupings established above. The 95th percentile of the distribution was determined for the five metrics that decrease in value with increasing anthropogenic impact: Total Taxa Richness; EPT Taxa Richness; Beck's Index; Shannon Diversity; and Percent Sensitive Individuals. Since the Hilsenhoff Biotic Index metric increases in value with increasing anthropogenic impact, the 5th percentile of the distribution was determined for this metric. Some metrics showed variability in the 95th or 5th percentiles with drainage area (Table 19, Figure 24).

Table 19. 95th (5th for Hilsenhoff Biotic Index) percentiles of all samples for each core metric by drainage area range.

Metric	**Drainage area range** (square miles)						
	0 to 3	**3 to 10**	**10 to 25**	**25 to 50**	**50 to 100**	**100 to 500**	**500 to 1,000**
	95th (5th for Hilsenhoff Biotic Index) percentiles of all samples by drainage area range						
Total Taxa Richness	32	34	35	34	34	33	33
EPT Taxa Richness (PTV 0-4 only)	18	20	20	20	19	19	17
Beck's Index (version 3)	37	41	37	35	31	28	22
Hilsenhoff Biotic Index	1.69	1.78	2.03	2.15	2.55	3.09	3.10
Shannon Diversity	2.83	2.90	2.90	2.88	2.87	2.87	2.95
% Sensitive Individuals (PTV 0-3 only)	88.0	84.1	82.6	81.0	78.5	65.5	68.3

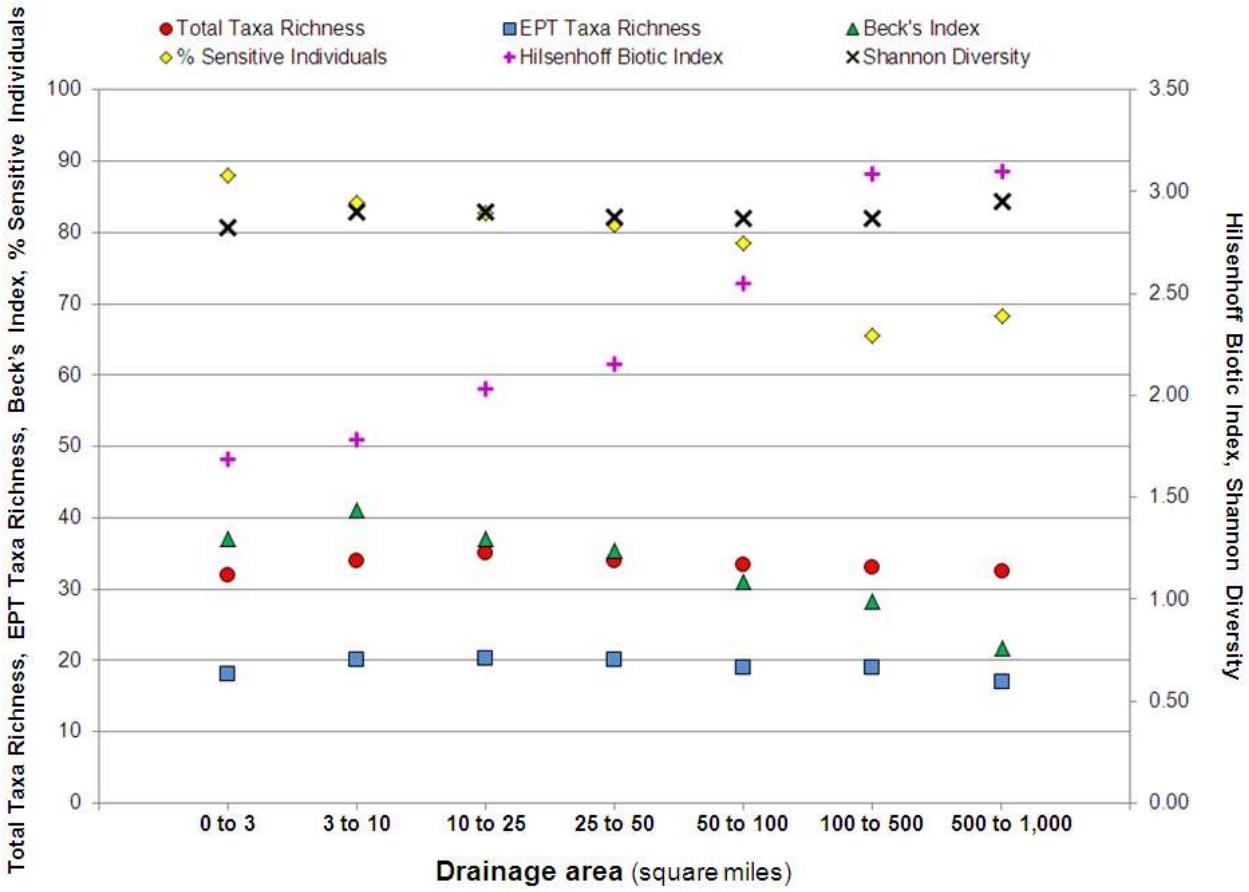


Figure 24. Plot of 95th (5th for Hilsenhoff Biotic Index) percentiles of all samples for each core metric by drainage area range.

In order to incorporate the variability of metric scores with drainage area in setting biological expectations through metric standardization values, PADEP decided to set two sets of standardization values for each selected core metric (Table 20). One set of metric standardization values applies to smaller streams – generally first through third order streams draining less than 25 square miles of land. The other set of metric standardization values applies to larger streams – generally fifth order and larger streams draining more than 50 square miles of land. The metric standardization values were chosen based on the 95th and 5th percentile values of the distributions. For larger streams, consideration was also given to the distribution of metric values for samples from streams larger than 1,000 square miles.

Table 20. Metric standardization values.

Metric	Metric Standardization Values	
	smaller streams most 1st to 3rd order < 25 square miles	**larger streams** most 5th order and larger > 50 square miles
Total Taxa Richness	33	31
EPT Taxa Richness (PTV 0-4 only)	19	16
Beck's Index (version 3)	38	22
Hilsenhoff Biotic Index	1.89	3.05
Shannon Diversity	2.86	2.86
% Sensitive Individuals (PTV 0-3 only)	84.5	66.7

To calculate the index of biological integrity, observed metric values are first standardized using the standardization values (Table 20) and the following standardization equations.

The Hilsenhoff Biotic Index metric values are expected to increase in value with increasing anthropogenic stress and are standardized using the following equation:

Hilsenhoff Biotic Index standardized score =
(10 – observed value) / (10 – standardization value) * 100

The other five core metrics values are expected to decrease in value with increasing anthropogenic stress and are standardized using the following equation:

Standardized metric score = observed value / standardization value * 100

Once the observed metric values are standardized, the standardized metric scores are adjusted to maximum value of 100 if necessary. Detailed examples of metric calculation and standardization along with index calculation are presented in Appendix C. By standardizing metrics and setting a maximum value of 100 for the standardized metrics, the resulting adjusted standardized metric scores can range from maximum values of 100 to minimum values of zero, with scores closer to zero corresponding to increasing deviation from the expected reference condition and progressively higher values corresponding more closely to the biological reference condition (Barbour et al. 1995). This approach establishes upper bounds on the expected condition and moderates effects of metrics that may respond in some manner other than a monotonic response to stress. The index of biological integrity is calculated by calculating the arithmetic mean of these adjusted standardized metric values for the six core metrics, resulting in a multimetric index of biological integrity score that can range from 0 to 100. To get a score of zero, a sample would have to contain no organisms at all.

In order to incorporate the variability of metric scores with annual seasons in setting biological expectations, PADEP chose to implement different use attainment benchmarks as discussed below rather than adjust metric standardization values.

INDEX PERFORMANCE EVALUATION

Biological Condition Discrimination

Across stream sizes, the IBI exhibited excellent ability to measure gradients of anthropogenic disturbance as defined by the condition categories in both the November to May time frame (Figure 25) and the June to September time frame (Figure 26).

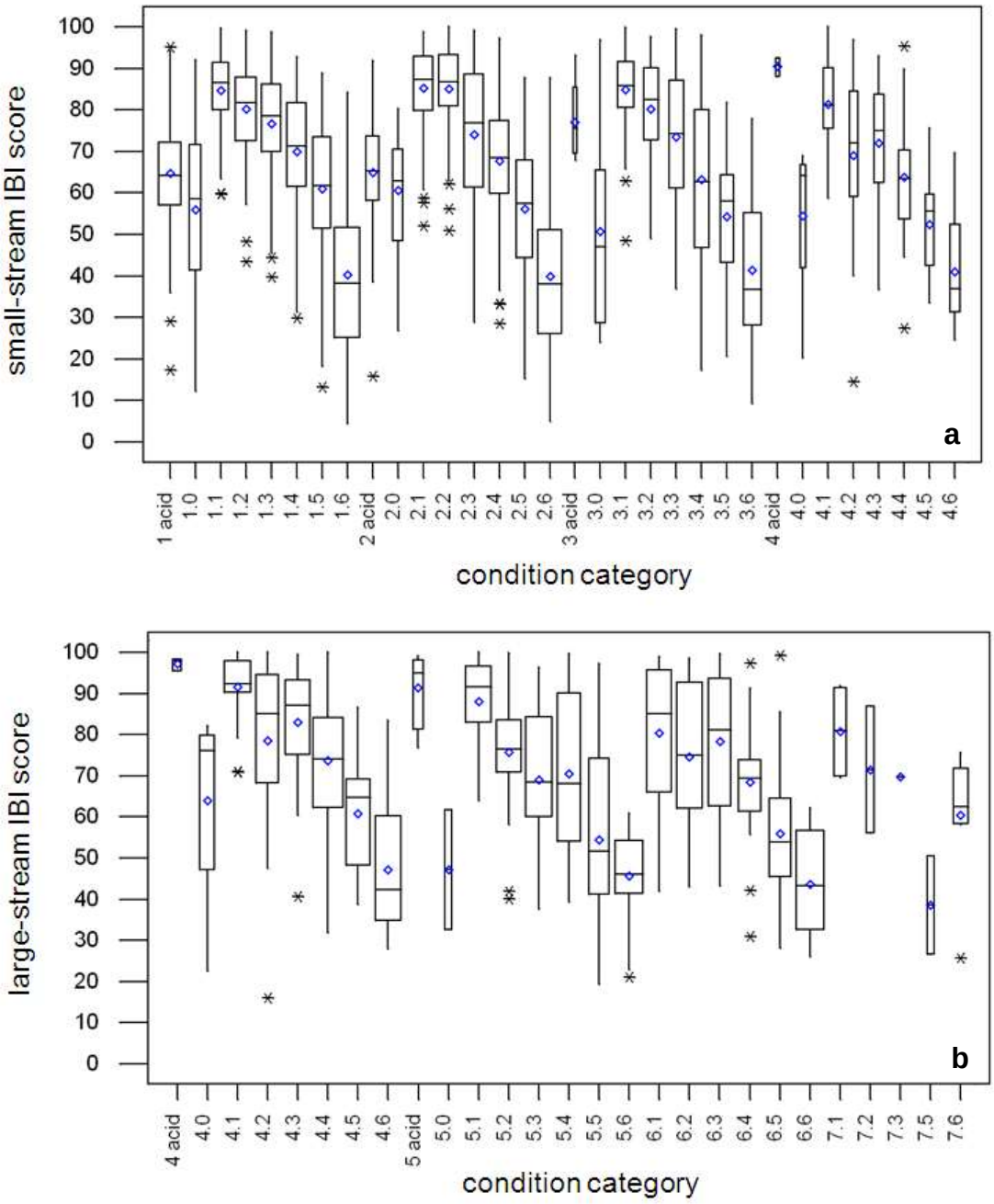


Figure 25. Boxplots of (a) small-stream IBI scores and (b) large-stream IBI scores by condition category for the November to May time frame. Within each figure, box widths are proportional to number of samples in each category. Total numbers of samples in each figure are: 2,110 samples for (a) and 473 samples for (b).

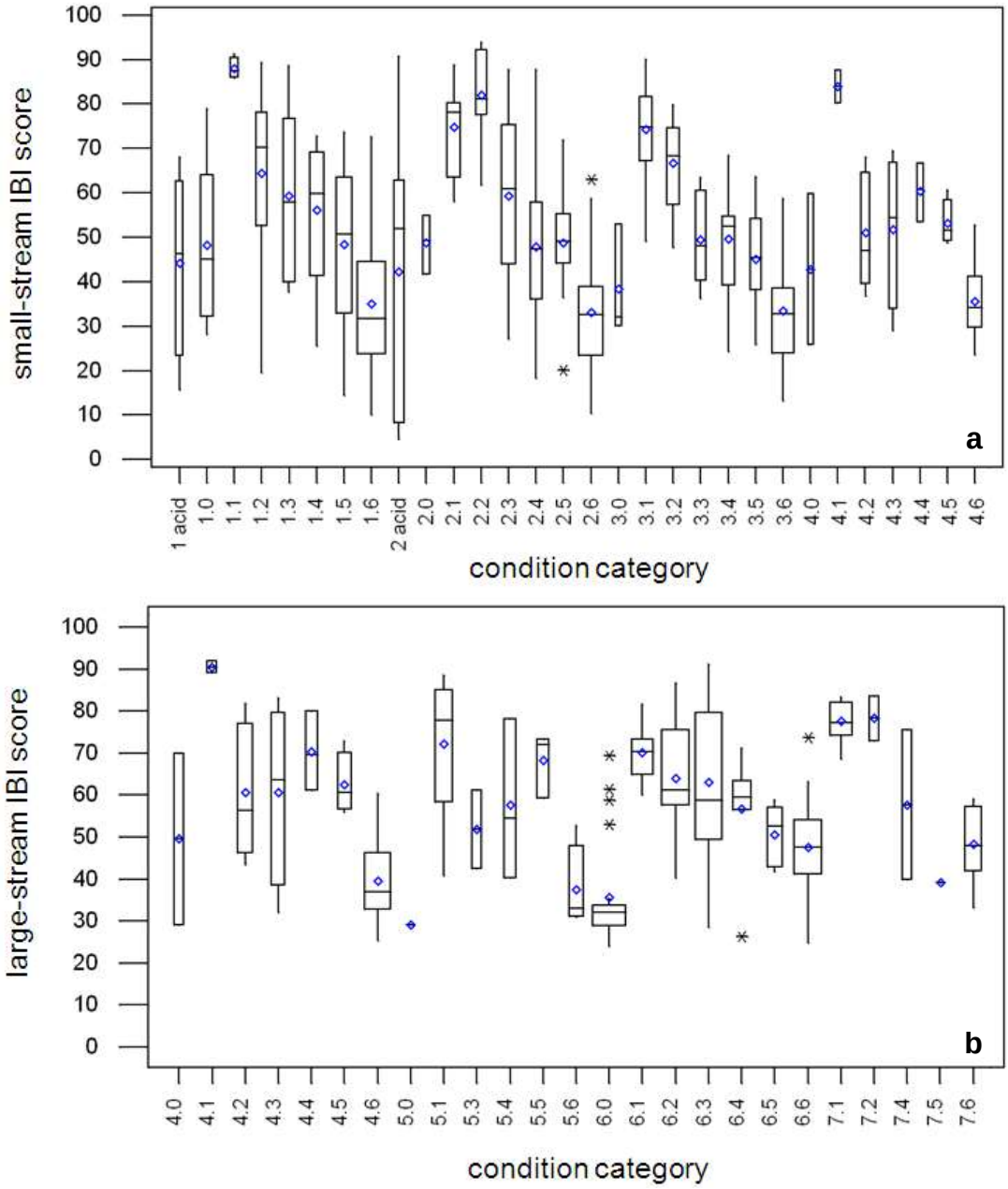


Figure 26. Boxplots of (a) small-stream IBI scores and (b) large-stream IBI scores by condition category for the June to September time frame. Within each figure, box widths are proportional to number of samples in each category. Total numbers of samples in each figure are: 353 samples for (a) and 169 samples for (b).

Using the same calculations as used for negative-response metrics above, discrimination efficiencies were calculated for the small-stream IBI and large-stream IBI across stream size ranges and time of year (Table 21). These IBI discrimination efficiencies further support the excellent ability of the IBIs to distinguish between reference conditions and severely impacted conditions.

Table 21. Discrimination efficiencies of the small-stream IBI and large-stream IBI by drainage area range and by seasons.

IBI	Season	Drainage area range (square miles)						
		0 to 3	3 to 10	10 to 25	25 to 50	50 to 100	100 to 500	500 to 1,000
		Discrimination Efficiency						
small-stream IBI	Nov - May	97%	98%	100%	100%			
	Jun - Sep	100%	100%	100%	100%			
large-stream IBI	Nov - May				100%	100%	100%	75%**
	Jun - Sep				100%	100%	94%	100%

** there were only eight "condition 6" samples from sites draining 500 to 1,000 square miles collected in the November to May time frame

The ability of the IBI to quantifiably differentiate biological communities among the condition categories as defined for this project strongly supports the utility of the IBI in measuring the biological condition of benthic macroinvertebrate communities in Pennsylvania's wadeable, freestone, riffle-run streams.

Intrasite Spatial Variability

Duplicate biological samples were taken at 56 sites and triplicate samples were taken at one site – each replicate set collected on the same day within the same 100-meter reach of stream. Analysis of all the replicate samples can provide an estimate of IBI intrasite spatial precision. These estimates of IBI and metric precision incorporate natural intrasite spatial variability and methodological variability.

Results of an analysis of variance (ANOVA) on the intrasite, same-day replicated sample data with site as a factor provides an estimate of variation for each set of replicated samples (Table 22). For purposes of this analysis, the small-stream IBI was applied to samples from streams draining less than 50 square miles and the large-stream IBI was applied to samples from streams draining more than 50 square miles. The metric values used in the ANOVA procedures were standardized and adjusted as described above and in Appendix C so that the relative magnitudes would be comparable with the IBI scores. The ANOVA mean square error (MSE) provides an estimate of within site standard deviation and can be used to calculate confidence intervals around a score. The lower the standard deviation, as estimated by the ANOVA MSE, the more confident we can be in methodological precision at a given site. The one-tailed 90% confidence intervals were calculated according to the following equation:

$$\text{One-tailed 90\% Confidence Interval} = 1.282 \times [(\text{ANOVA MSE})^{0.5} / (\text{number of samples})^{0.5}]$$

Table 22. Intrasite spatial precision estimates for IBI scores and each core metric based on ANOVA results. The ANOVA mean square error (MSE) estimates intrasite standard deviation. Coefficients of variation (CV) were calculated for each sample pair (or triplet) and then averaged across all sample pairs. "*s*" indicates standardized metric values. "*r*" indicates raw metric values. For simplicity, the small-stream procedures were applied to samples from sites draining less than 50 square miles and the large-stream procedures were applied to samples from sites draining more than 50 square miles. There were only two large-stream samples from one site in the November to May time frame, so ANOVA was not possible: standard deviations and CVs are reported.

Metric		small-stream						large-stream				
		November to May 79 samples from 39 sites			June to September 8 samples from 4 sites			November to May 22 samples from 11 sites			June to September 2 samples from 1 site	
		ANOVA MSE	90% CI (1 sample)	CV	ANOVA MSE	90% CI (1 sample)	CV	ANOVA MSE	90% CI (1 sample)	CV	standard deviation	CV
IBI score		16.2	**5.16**	**5.7%**	21.1	**5.89**	**16.6%**	15.2	**5.00**	**5.9%**	**0.00**	**0.0%**
Total Taxa Richness	*s*	50.5	9.11	8.9%	109.0	13.38	22.4%	84.2	11.76	8.4%	0	0.0%
	r	5.8	3.09	9.5%	11.9	4.42	22.4%	8.1	3.65	8.4%	0	0.0%
EPT Taxa Richness (PTV 0-4 only)	*s*	75.5	11.14	22.8%	31.2	7.16	51.1%	130.0	14.62	18.4%	0	0.0%
	r	2.8	2.14	24.0%	1.1	1.36	51.1%	3.6	2.45	18.4%	0	0.0%
Beck's Index (version 3)	*s*	63.4	10.21	29.9%	4.3	2.67	14.5%	85.5	11.85	24.8%	0	0.0%
	r	11.2	4.29	33.6%	0.6	1.01	14.5%	9.6	3.98	25.4%	0	0.0%
Hilsenhoff Biotic Index	*s*	11.8	4.40	2.6%	0.7	1.09	1.5%	15.1	4.98	5.1%	0.03	0.7%
	r	0.1	0.39	7.0%	0.0	0.09	1.0%	0.1	0.36	4.9%	0.45	0.6%
Shannon Diversity	*s*	33.4	7.41	5.9%	121.0	14.10	20.4%	15.0	4.97	4.8%	0.00	0.1%
	r	0.0	0.21	6.0%	0.1	0.40	20.4%	0.0	0.14	4.8%	0.05	0.1%
% Sensitive Individuals (PTV 0-3 only)	*s*	60.7	9.99	16.3%	10.3	4.11	60.2%	28.4	6.83	21.7%	0.24	3.6%
	r	44.4	8.54	18.2%	7.4	3.48	60.2%	12.7	4.57	21.7%	0.36	3.6%

The results of the intrasite spatial precision estimate analysis (Table 22) suggest that the Hilsenhoff Biotic Index metric – standardized and adjusted or raw – tends to vary less intrasite than other core metric and the IBI scores. The IBI scores tend to vary relatively little intrasite compared with the standardized and adjusted metric values. These results highlight that the IBI, by combining the six metrics into a single index, attenuates much of the intrasite variability of each metric individually, providing a more spatially stable indication of biotic condition than any one metric could alone.

Temporal Variability

Two-hundred ninety-two sites were sampled on more than one date, ranging from two to twelve samples taken over time at a given site, for a total of 813 samples. Analysis of all samples from the same sites over time can provide an estimate of temporal variability of the IBI and metric scores. As with the intrasite spatial precision estimates, the estimates of IBI and metric temporal precision incorporate natural intrasite spatial variability and methodological variability, but they also incorporate natural temporal variability and variability due to changes in condition over time.

The same approach was used to analyze temporal variability as was used to evaluate intrasite spatial variability described above. The results of the temporal precision estimate analysis (Table 23) suggest that the Percent Sensitive Individuals metric tends to vary the most of any of the metrics or the IBI score over time at a site. Like the intrasite precision estimates, the temporal precision results also highlight that the IBI attenuates much of the temporal variability of each metric individually, providing a more temporally stable indication of biotic condition than any one metric could alone.

In the temporal precision estimate dataset, there was not any substantial relationship between stream size and variability of IBI scores at sites over time as measured by standard deviation of IBI scores (Figure 27).

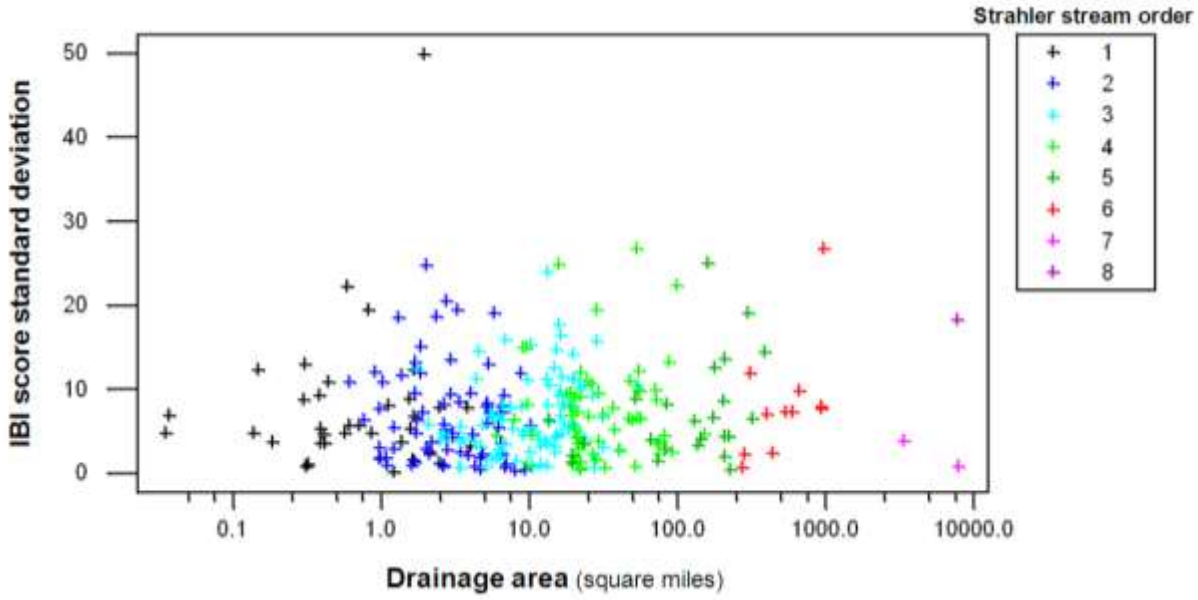


Figure 27. Standard deviation of IBI scores of samples at sites in the temporal precsion estimate dataset by drainage area.

Table 23. Temporal precision estimates for IBI scores and core metrics based on ANOVA results. The ANOVA mean square error (MSE) estimates intrasite standard deviation. Coefficients of variation (CV) were calculated for each sample pair (or triplet or quadruplet…) and then averaged across all sample pairs. "*s*" indicates standardized metric values. "*r*" indicates raw metric values. For simplicity, the small-stream procedures were applied to samples from sites draining less than 50 square miles and the large-stream procedures were applied to samples from sites draining more than 50 square miles.

Metric		small-stream						large-stream					
		November to May 384 samples from 137 sites			June to September 26 samples from 12 sites			November to May 78 samples from 26 sites			June to September 26 samples from 7 sites		
		ANOVA MSE	**90% CI** (1 sample)	**CV**	**ANOVA MSE**	**90% CI** (1 sample)	**CV**	**ANOVA MSE**	**90% CI** (1 sample)	**CV**	**ANOVA MSE**	**90% CI** (1 sample)	**CV**
IBI score		48.9	**8.96**	**8.8%**	95.7	**12.54**	**19.6%**	69.0	**10.65**	**10.3%**	18.5	**5.51**	**4.8%**
Total Taxa Richness	*s*	115.0	13.75	10.9%	101.0	12.88	13.3%	128.0	14.50	12.5%	103.0	13.01	10.0%
	r	16.6	5.22	13.2%	16.1	5.14	14.8%	15.5	5.05	13.2%	12.1	4.46	11.3%
EPT Taxa Richness (PTV 0-4 only)	*s*	138.0	15.06	18.5%	89.5	12.13	23.8%	185.0	17.44	17.3%	78.8	11.38	10.7%
	r	6.3	3.21	19.7%	4.8	2.81	24.7%	7.9	3.59	20.8%	2.0	1.82	10.7%
Beck's Index (version 3)	*s*	127.0	14.45	22.8%	94.4	12.46	36.9%	132.0	14.73	14.2%	142.0	15.28	24.6%
	r	21.9	6.00	23.7%	17.9	5.42	37.5%	16.0	5.13	19.7%	10.4	4.13	26.4%
Hilsenhoff Biotic Index	*s*	53.1	9.34	7.3%	222.0	19.10	22.6%	71.3	10.83	8.3%	18.5	5.51	4.5%
	r	0.4	0.79	15.6%	1.5	1.57	21.2%	0.4	0.81	15.4%	0.1	0.38	6.1%
Shannon Diversity	*s*	96.1	12.57	10.1%	131.0	14.67	14.1%	120.0	14.04	10.5%	33.5	7.42	5.3%
	r	0.1	0.38	10.7%	0.1	0.45	14.4%	0.1	0.42	10.8%	0.0	0.24	5.7%
% Sensitive Individuals (PTV 0-3 only)	*s*	215.0	18.80	23.6%	361.0	24.36	65.7%	337.0	23.53	27.7%	133.0	14.78	16.5%
	r	157.0	16.06	23.8%	258.0	20.59	65.7%	197.0	23.53	30.2%	59.1	9.86	16.5%

Application to an Independent Dataset

To further evaluate its performance, the IBI was applied to 116 samples collected from 112 wadeable, freestone, riffle-run stream sties in Pennsylvania – using the same sampling collection and processing protocol outlined above – for the Regional Environmental Monitoring and Assessment Program (REMAP), a project coordinated by the United States Environmental Protection Agency. None of the REMAP samples were used in the IBI development process.

Samples for REMAP were collected at sites across the state (Figure 28) between March 30, 2005 and May 27, 2005 (Figure 29) by biologists with the United States Environmental Protection Agency, the Delaware River Basin Commission, or SoBran, a private contractor.

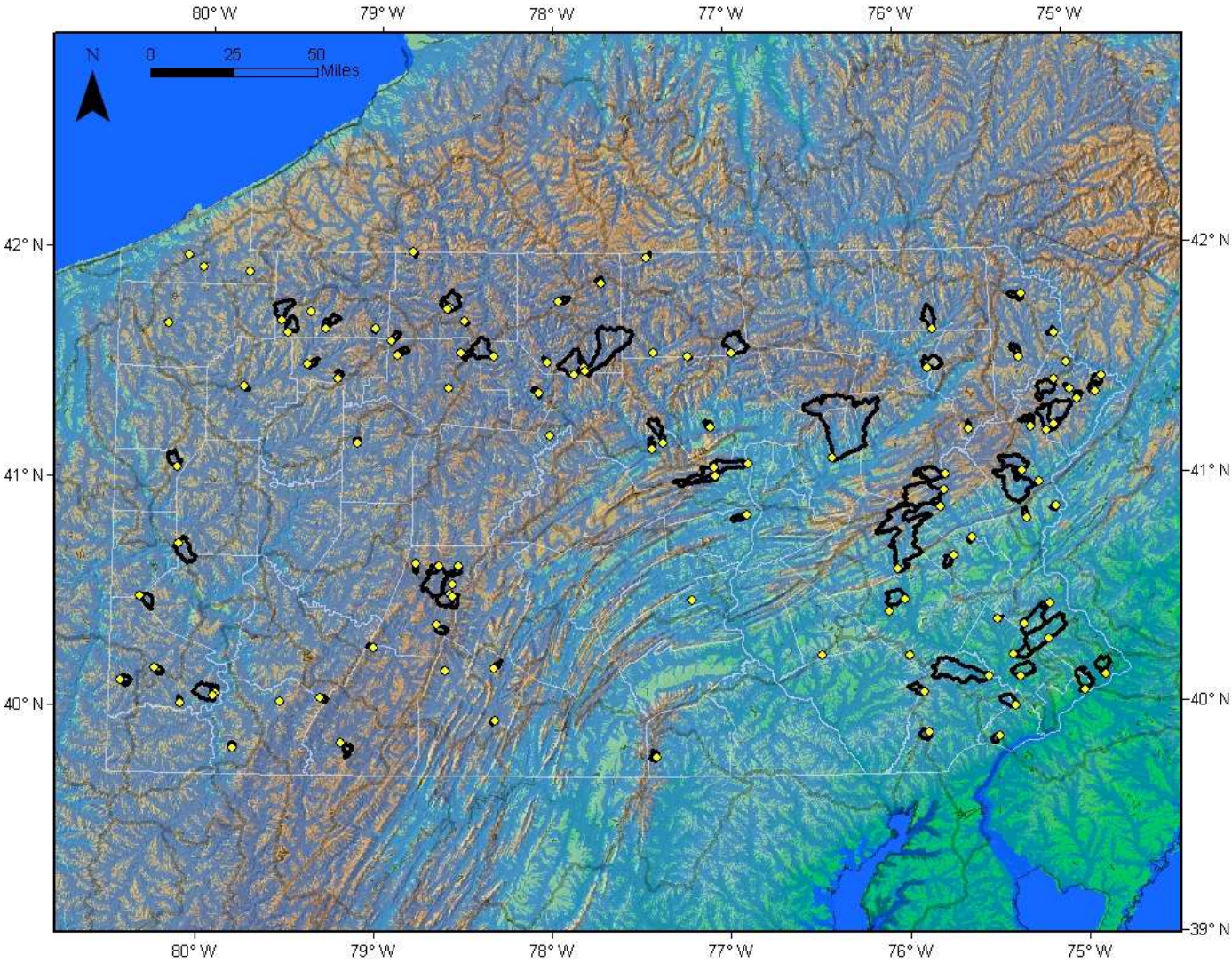


Figure 28. Map of 112 sites and associated basins for the 116 samples collected in Pennsylvania for the REMAP project coordinated by the United States Environmental Protection Agency.

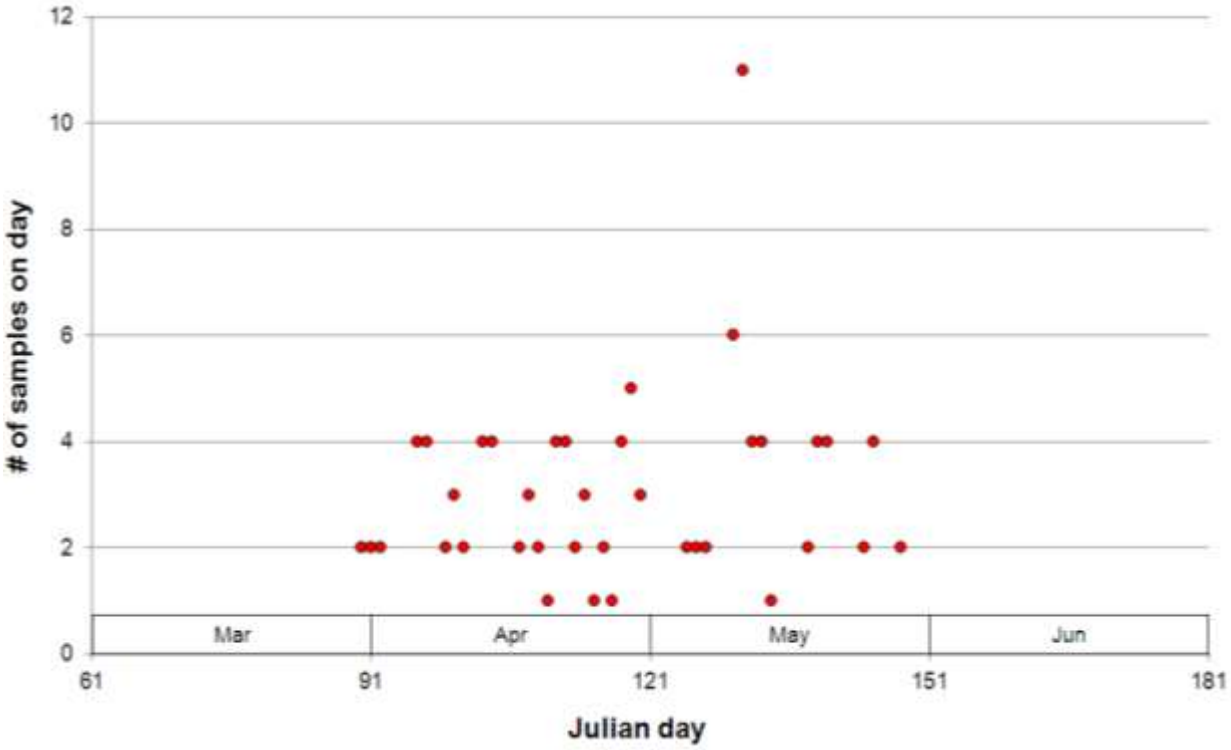


Figure 29. Distribution of REMAP samples by Julian day. All REMAP samples were collected in the spring of 2005.

Most REMAP samples were collected from sites on relatively small, first and second Strahler order streams draining less than ten square miles of land (Table 24).

Table 24. Number of REMAP samples by drainage area range and Strahler stream order.

Drainage area range (square miles)	**Strahler stream order**				
	1	2	3	4	5
0 to 3	27	12			
3 to 10	11	26			
10 to 25		10	11		
25 to 50			9		
50 to 100			2	6	
100 to 500				1	1

As with the IBI development dataset, the highest gradient streams tended to be smaller streams (Figure 30), with less relationship between slope and elevation (Figure 31), and with larger sites being at mostly lower elevations (Figure 32). Since the REMAP sampling was conducted in a two-month window, from March to May in 2005, seasonal considerations (Figure 33) are as not much of a concern with the REMAP dataset as with the IBI development dataset.

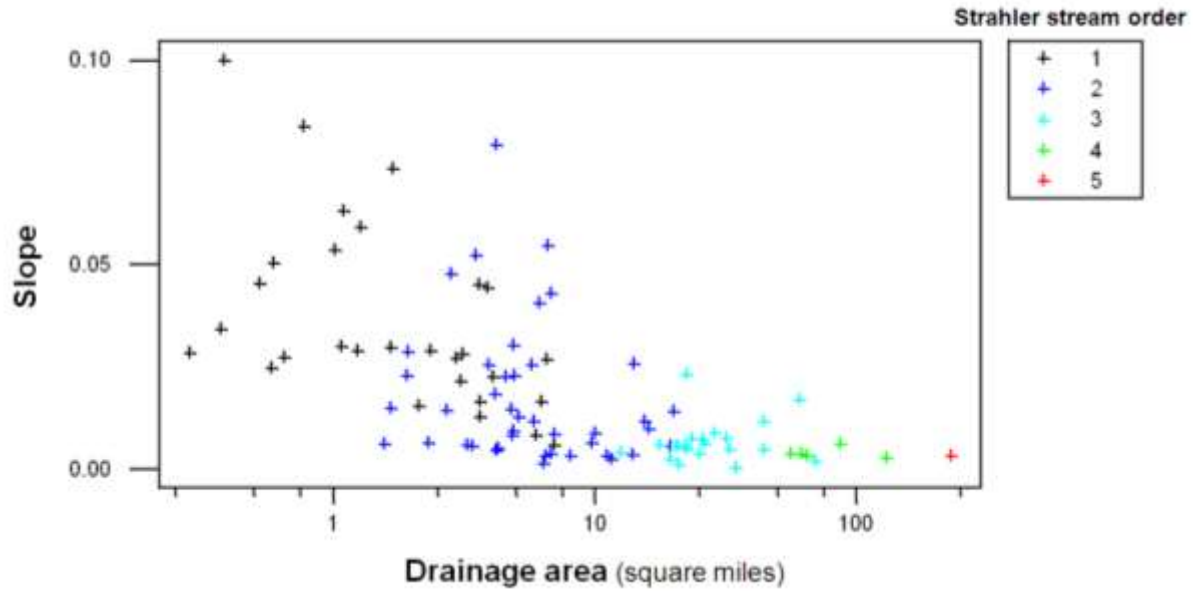


Figure 30. Relationship of slope and drainage area at 98 REMAP sites for which slope data was available. Note logarithmic scale for drainage area.

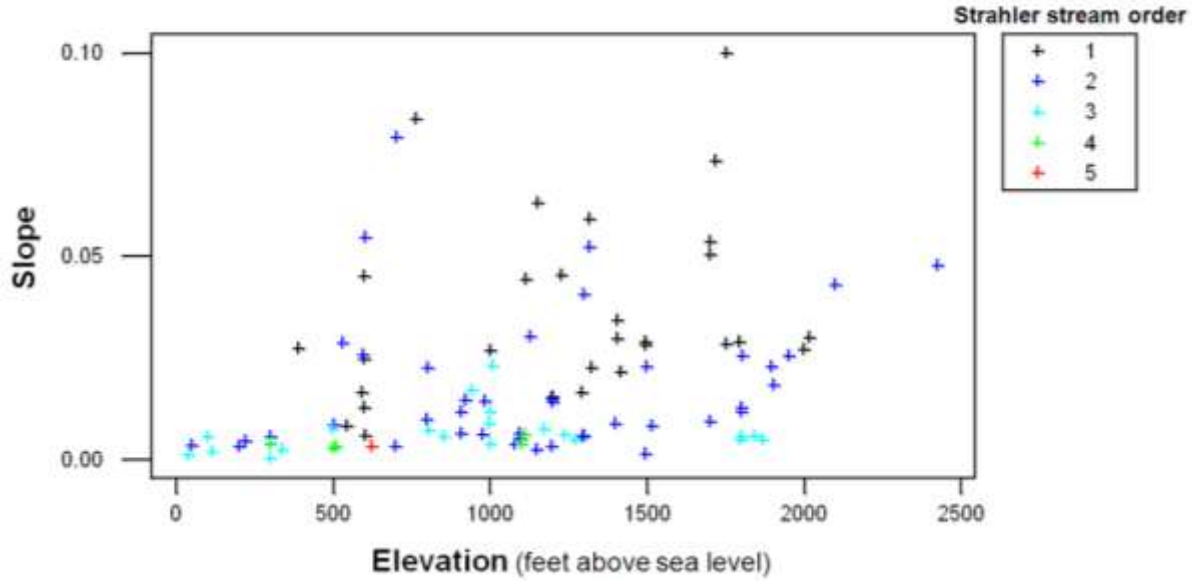


Figure 31. Relationship of slope and elevation at 98 REMAP sites for which slope data was available.

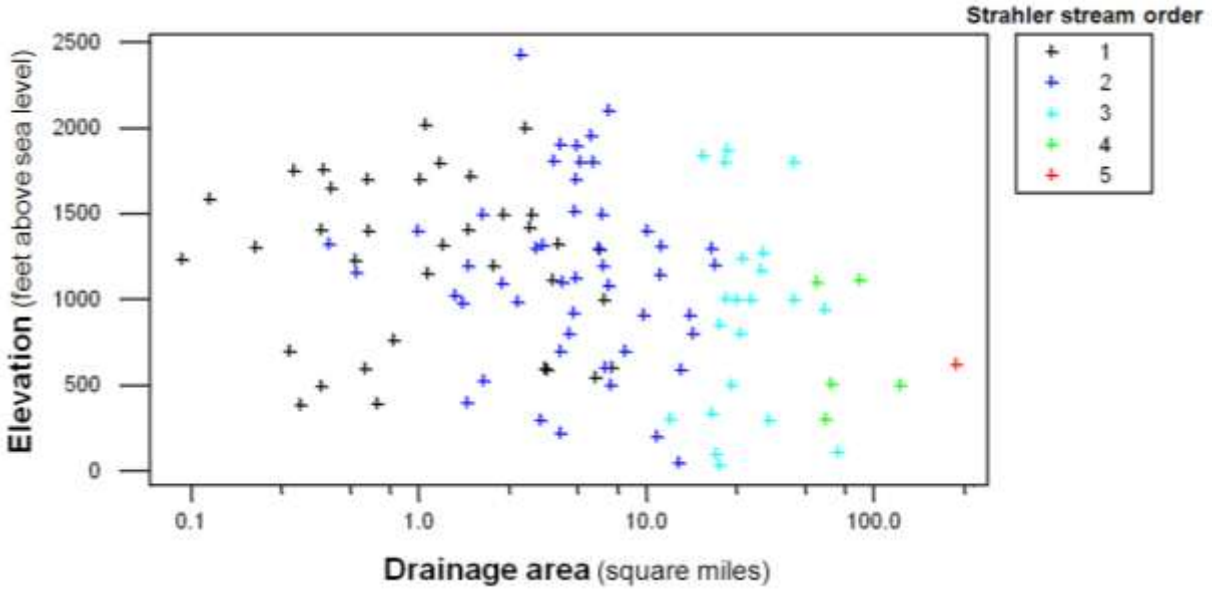


Figure 32. Relationship of elevation and drainage area at 112 REMAP sites. Note logarithmic scale for drainage area.

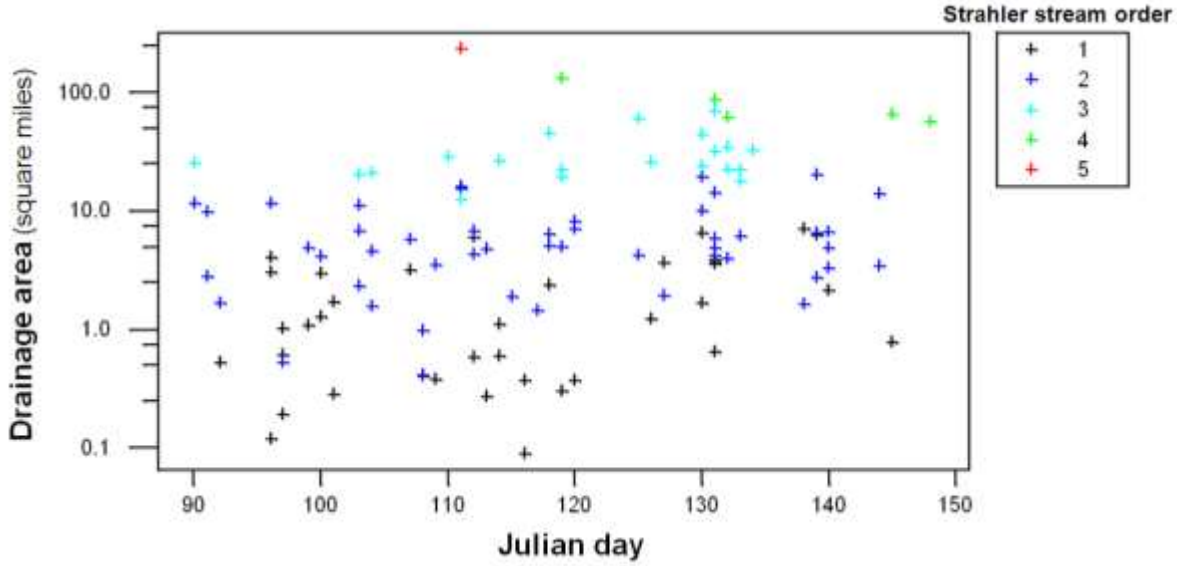


Figure 33. Relationship of drainage area and Julian day of sample collection for 116 REMAP samples. Note logarithmic scale for drainage area.

For purposes of comparison with the IBI development dataset, the abiotic condition determination process described above – without biotic screening for acid deposition impacts – was applied to the REMAP sites. Condition index scores for the REMAP sites ranged from 198 to -180. Since the REMAP dataset consisted of many fewer samples than the IBI development dataset, the REMAP samples were simply divided into five groups based on condition index scores in order to assess the efficacy of the IBI in distinguishing among sites variously impacted by human activities. The large-stream IBI was applied to REMAP samples from sites draining more than 50 square miles and the small-stream IBI was applied to REMAP samples from sites draining less than 50 square miles. Applied in this way to the REMAP samples, the IBI displayed marvelous ability to distinguish among various levels of human impacts as measured by the condition index (Figure 34).

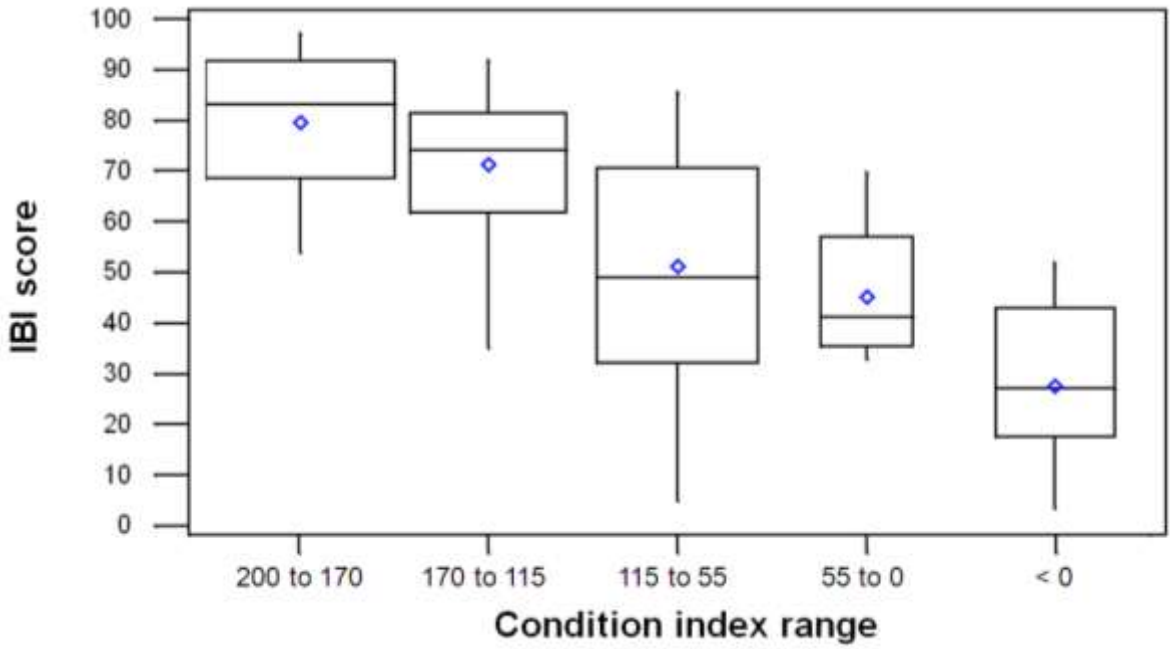


Figure 34. Boxplot of IBI scores and condition index ranges for 116 REMAP samples. The large-stream IBI was applied to samples from sites draining more than 50 square miles. The small-stream IBI was applied to samples from sites draining less than 50 square miles.

Ten of the REMAP sub-samples contained more than 240 organisms and eight of the REMAP sub-samples contained less than 160 organisms. Whether these samples with sub-samples out of the target range or organisms are included in the REMAP analyses or not, the IBI demonstrates very good ability to distinguish among sites variously impacted by human activities.

Duplicate biological samples were taken at four REMAP sites on the same day. The variability of the IBI scores for each pair of duplicate REMAP samples was very low, with the standard deviation of duplicate sample pairs ranging from 0.49 for McMichaels Creek to 3.46 for O'Donnell Creek (Table 25). If we run these four duplicate sample pairs through a one-way ANOVA with IBI score as the response and site as the factor and apply the confidence interval calculation discussed above, we get a one-sample 90% confidence interval of 3.47 IBI points.

Table 25. IBI scores for REMAP sites sampled twice on the same day.

Stream Name	Sampling date	Drainage area (square miles)	Condition Index	IBI scores		IBI score standard deviation
				small-stream	large-stream	
Bush Kill	May 27, 2005	55.9	185		**89.1**	2.55
					92.7	
McMichaels Creek	May 24, 2005	64.7	127		**82.9**	0.49
					83.6	
O'Donnell Creek		0.8	182	**70.9**		3.46
				75.8		
Sandy Run	May 12, 2005	22.4	77	**46.7**		3.25
				51.3		

Large Wadeable Rivers

The preceding analysis only considered samples from streams draining less than 1,000 square miles of land. However, there were 29 samples collected and processed using the same methodology described above from 26 sites draining more than 1,000 square miles of land. The sample collection and processing methods described above are intended to be applied to *wadeable*, freestone, riffle-run streams. These methods focus on sampling riffle areas because these habitats are typically the most productive in riffle-run streams. The substrate area sampled and the target number of organisms sub-sampled with these methods yield sufficient representation of the benthic communities in these streams to assess biological integrity and anthropogenic impacts with reasonable accuracy and precision.

Although PADEP does not currently have strict guidelines for determining the upper limit of stream size for which these methods are tenable, there is general recognition that these methods may not sufficiently represent the benthic or overall biotic communities in the largest of Pennsylvania's streams and rivers. As a river system becomes larger and larger, these sampling methods – which observe a fixed area of one habitat type regardless of the size of the stream or the proportion of various habitat types in the stream – represent smaller and smaller proportions of the whole stream benthos and biota. PADEP is currently working to develop methods to assess ALUs in larger streams and rivers, including non-wadeable, dam-pool rivers like the lower Monongahela River, the lower Allegheny River, and the Ohio River. These methods will likely include sampling and assessment of various biotic assemblages (e.g., benthic macroinvertebrates, fish, mussels) and may utilize different sampling equipment (e.g., Hester-Dendy type multiplate samplers) and approaches (e.g., littoral sampling, sampling larger areas, targeting different habitat types, identifying chironomids to finer taxonomic levels) to evaluate the benthos in larger streams and rivers.

As noted above, PADEP does not have hard-and-fast guidelines for determining the upper limit of stream size to which the methods outlined in this report can be tenably applied. Some sections of some streams are obviously and always wadeable throughout their course (e.g., first-order headwater creeks). In larger streams and rivers, some areas (e.g., shallow riffles) are consistently wadeable while other areas (e.g., deeper pools and runs) may never be wadeable or may only be wadeable during low flow conditions. In certain situations it should be clear that these methods do not apply (e.g., if a stream is not wadeable in over 90% or more of its channel area under base flow conditions). If a stream is only unwadeable in one small spot of one deep pool in the sampling reach at baseflow, but wadeable throughout the rest of the reach, it is likely tenable to apply these sampling methods. If a stream is unwadeable only during 100-year flood flows, but entirely wadeable during other flows, it is likely tenable to apply these methods. Between these extremes, discretion must be used in applying the sampling and assessment methods outlined in this report to the largest of wadeable streams and rivers. In rivers where riffle habitat represents exceedingly small proportions of the overall channel area, the ALU assessment methods presented in this report should be not applied.

With that in mind, the large-stream IBI developed in this project was applied to the 27 samples from 24 sites draining more than 1,000 square miles of land in this dataset, and it performed well even with these large river samples (Table 26, Figure 35). Of these large river samples, the four samples that scored highest on the IBI were from the highly forested

upper Delaware River, at locations that had total habitat scores between 180 and 195. Two samples from the middle Delaware River – 10 to 15 miles upstream of the confluence with Lehigh River – scored noticeably lower on the IBI than samples from the upper river. Although the middle Delaware basin is still mostly forested, there are more anthropogenic impacts here than in the upper reaches along with natural changes that occur as the stream flows downstream. A sample from the lower Delaware River – near Trenton, New Jersey – scores even lower than samples from the middle part of the river. Note that the samples from the upper Delaware River were collected in April while the samples from the middle and lower parts of the river were collected in August and September, so we may be seeing some drop in IBI scores related to sampling season, although this is difficult to determine conclusively with such a small number of samples at these sites. We must also be mindful that the flow and thermal patterns in the Delaware River are hugely influenced by releases from upstream drinking water reservoirs.

A sample from Sinnemahoning Creek – a highly forested basin – collected early September 2007 scored 57.6 on the IBI. We might expect the score from such a highly forested basin to be higher. This sample was collected in early September, so we are likely seeing some drop in the IBI score due to sampling season. The physical habitat may also be naturally limiting the macroinvertebrate community at this location. A few in-stream habitat parameters were scored quite low here, which may reflect predominance of bedrock substrate which is not uncommon throughout the lower reaches of Sinnemahoning Creek.

Two samples from the same location on the lower Juniata River scored fairly high on the IBI, with a sample collected in August 2007 scoring about five points higher than a sample collected mid-October 2003. The Juniata River basin encompasses a variety of human impacts including mine drainage in some upper reaches and some population centers, with the most prevalent impact being agriculture. This basin also drains a fair amount of calcareous geologies.

A sample from the mouth of French Creek – near Franklin – collected mid-September 2007 scored 70.1 on the IBI. There is a fair amount of agriculture and low-density residential land use in this basin, but the benthic macroinvertebrate community appears to be in relatively good condition.

Two samples from a site on the middle Allegheny River – near Parker, about a mile downstream of the confluence with Clarion River – score about 25 points differently on the IBI, with a sample from early May 2003 scoring 64.6 and a sample from mid-October 2001 scoring 38.8. It appears seasonal considerations may explain much of this large difference in IBI scores with the May sample containing much higher mayfly diversity and abundance as well as higher stonefly and caddisfly diversity. Beetle diversity was much higher in the May sample as well. However, with such a small dataset from large rivers, it is difficult to determine conclusively what factors contribute to variability in sampled taxa. Some of the large differences we see in IBI scores at some these sites over time may have as much to do with considerations of patchy habitat and organismal distributions in larger systems as it does with the seasonal patterns we see in smaller systems.

Samples from various locations along the Susquehanna River score in the middle of the pack on the IBI among large river samples. We have two samples from the middle reaches of the “North Branch” Susquehanna River – near Towanda, about a mile upstream of the

confluence with Towanda Creek – collected early October 2003 and mid-July 2008 that score 58.6 and 59.7 on the IBI, respectively. These two samples score remarkably close on most core metrics and had fairly similar taxa lists. There was one sample from further downstream on the "North Branch" – about four miles west of Nanticoke, between the confluences of Hunlock Creek and Shickshinny Creek – collected mid-October 2007 that scored 47.3 on the IBI. This lower IBI score compared to upstream is mostly attributable to the Hilsenhoff Biotic Index and Percent Sensitive Individuals metrics, reflecting the relatively lower abundance of mayflies compared to two samples from further upstream as well as relatively high abundances of a few high-PTV snail families and Stenelmis beetles at the site near Nanticoke compared to the site near Towanda. On the West Branch Susquehanna, we have one sample collected early September near Jersey Shore between the confluences of Antes Creek and Larrys Creek that scores 53.9 on the IBI. In the lower reaches of the Susquehanna River, we have two samples: one sample collected mid-October near Sunbury just downstream of the confluence of the West Branch and "North Branch" upstream of the confluence with Shamokin Creek that scores 57.7 on the IBI; and one sample collected early October near Wrightsville and Columbia between the confluence of Chiques Creek and Kreutz Creek that scores 61.5 on the IBI.

The large river samples that scored lowest on the IBI were collected in late April 2008 from the lower Lehigh River at locations downstream of or within the highly urbanized areas of Allentown, Bethlehem, and Easton. The lower Lehigh River basin is also impacted by agriculture and anthracite coal mine drainage as well as some calcareous geologies. However, a sample from the lower Lehigh River collected mid-August 2007 scored over 20 points higher on the IBI than a sample collected less than a half mile upstream in April 2008. This suggests the possibility of a counterintuitive seasonal pattern to the taxa (mayfly diversity and abundance were notably higher in the August 2007 sample than the April 2008 sample), an intervening pollution event, anomalous weather events, and/or patchiness of habitat at this location.

Other samples that scored low on the IBI from heavily-impacted large rivers include samples from Mahoning River (early October 2007), Conemaugh River (late September 2007), Schuylkill River (early August 2007), and Youghiogheny River (late September 2007). These low IBI scores may be attributable in part to the samples being collected in late summer and early autumn, but it is likely that the substantial human impacts to these basins and rivers also drives down the IBI scores. On the Schuylkill River, the sample from further upstream – near Pottstown, just downstream of the confluence with Manatawny Creek – scores about 11 points higher on the IBI than the sample from further downstream – near western Philadelphia, just downstream of the confluence with Wissahickon Creek – where the urbanized land use is more intense. Note that both these Schuylkill River samples were collected a day apart in early August 2007.

It may be worth noting that the four upper Delaware River samples from April 2006 contain a number of taxa unique among these large river samples. These four samples were the only samples from sites draining more than 1,000 square miles that had more than two stonefly taxa. All four of these samples encountered Acroneuria and Perlesta stoneflies, with Paragnetina and Agnetina – fellow members of the Perlidae family – found in two and one of these samples, respectively. Three genera of Perlodidae stoneflies were also found in these samples: Cultus in three samples; Helopicus in two; and Isoperla in one. Acroneuria were only encountered in one other sample from a site draining more than 1,000

square miles: the mid-September sample from French Creek. Likewise, Perlesta and Perlodidae were only encountered in one other sample from a site draining more than 1,000 square miles: the early May sample from the Allegheny River. These four upper Delaware River samples accounted for the only records of Paragnetina among samples from sites draining more than 1,000 square miles. Cinygmula, Drunella, and Eurylophella mayflies were only found in these upper Delaware River samples as well, with Epeorus only being found in these samples as well as in one Lehigh River sample. In addition, among these large river samples, the only records of Rhyacophila and Lepidostoma caddisflies as well as Clinocera dance flies and Prosimulium blackflies are from the upper Delaware River samples. Although it is difficult to draw conclusions from such a small dataset it appears the upper Delaware River contains an unusual benthic community for such a large river.

Another interesting taxonomic phenomenon among these large river samples has to do with gastropods, which PADEP identifies to the family level. Hydrobiidae and Pleuroceridae snails are rarely seen in samples from smaller streams (usually only in smaller streams with substantial amounts of agriculture in their basin), but are somewhat common in samples from larger streams, especially those draining over 1,000 square miles.

Although the large-steam IBI appears to work fairly well when applied to this limited dataset of samples from large rivers (i.e., sites draining over 1,000 square miles), discretion must be used when applying this IBI to samples from such large rivers. The relatively small dataset of samples from such large rivers limits analysis of variability (i.e., estimates of spatial and temporal precision) in metric and IBI performance with samples from such large rivers.

As long as the area of riffle habitat relative to total channel area is not exceedingly low, the methods outlined in this project may be tenably applied to larger river systems. If riffle habitats represent a very small proportion of the total channel area in a larger river, these methods may be less appropriate to apply. The aquatic life uses of these lower gradient larger rivers may be better assessed by deploying different types of sampling equipment (e.g. Hester-Dendy multiplate samplers), targeting different habitats (e.g., pools, littoral areas), utilizing different levels of taxonomic identification (e.g., identifying chironomids to tribe, genus, or species), and incorporating assessments of other biological assemblages (e.g., fish, mussels, plankton, perphyton).

PADEP is currently working on assessment methods for non-wadeable rivers and larger partially-wadeable or sometimes-wadeable rivers. These developing methods may be better suited to conducting assessments in some of Pennsylvania's largest river systems, but – as an interim procedure – the methods outlined in this project can be tenably applied to larger systems with adequate riffle habitats that can be consistently and safely accessed by foot.

Table 26. Habitat scores, % land uses, IBI scores, and core metric values for the 27 samples from 24 sites draining more than 1,000 square miles of land.

Stream Name	Drainage area (square miles)	Month	Total Habitat Score	% land uses			IBI score large-stream	Core metric values					
				forest	agriculture	developed		Total Taxa Richness	EPT Taxa Richness (PTV 0-4)	Beck's Index (version 3)	Hilsenhoff Biotic Index	Shannon Diversity	% Sensitive Individuals (PTV 0-3)
Sinnemahoning Creek	1,033	9	167	93.5	3.0	0.2	**57.6**	21	10	7	4.76	2.46	14.6
Mahoning River	1,100	10	148	34.4	31.7	12.5	**27.8**	11	1	0	5.61	1.77	0.0
Schuylkill River	1,147	8	174	45.3	37.7	7.2	**48.6**	22	8	5	5.75	2.35	3.2
Lehigh River	1,227	4	153	59.2	18.2	6.6	**25.9**	12	2	3	5.83	0.86	0.5
French Creek	1,237	9		53.0	33.4	1.8	**70.1**	26	11	10	4.42	2.67	32.6
Lehigh River	1,357	4	185	55.5	20.5	7.9	**25.5**	13	1	2	5.84	0.91	2.8
Lehigh River	1,360	8	164	55.5	20.5	7.9	**49.9**	25	5	0	5.33	2.53	21.2
Conemaugh River	1,362	9	152	67.7	20.3	2.8	**32.9**	13	3	0	5.11	1.90	0.0
Delaware River	1,522	4	189	81.3	11.8	0.4	**79.1**	24	14	20	4.16	2.15	39.9
Delaware River	1,621	4	194	81.3	11.7	0.4	**94.7**	30	20	25	3.55	2.80	53.8
Delaware River	1,713	4	182	81.4	11.6	0.4	**87.6**	36	17	16	3.92	2.93	43.5
Youghiogheny River	1,713	9	153	65.4	23.0	3.0	**38.7**	15	5	0	4.45	1.70	9.2
Schuylkill River	1,879	8	173	39.8	38.2	10.2	**37.5**	13	5	2	5.27	1.92	5.0
Delaware River	1,906	4	190	79.9	13.0	0.4	**94.7**	33	19	24	3.52	2.96	49.8
Juniata River	3,352	8		69.5	21.9	2.3	**72.5**	30	11	7	4.02	2.84	34.6
		10					**67.0**	28	11	9	5.03	2.71	23.6
Delaware River	4,542	9		76.4	10.9	1.1	**54.9**	21	9	7	5.39	2.63	10.3
Delaware River	4,552	9		76.3	11.0	1.1	**68.7**	28	10	10	4.58	2.67	28.3
West Branch Susquehanna River	5,230	9	163	*land use not calculated for sites draining more than 5,000 square miles*			**53.9**	25	10	3	6.12	2.64	12.5
Delaware River	6,788	8					**48.9**	20	8	4	5.10	2.35	5.5
Allegheny River	7,663	5					**64.6**	34	11	7	5.91	3.05	18.8
		10					**38.8**	21	5	3	7.46	2.26	3.1
Susquehanna River	7,792	7	189				**59.7**	24	9	4	4.88	2.57	28.8
		10					**58.6**	24	10	6	4.92	2.60	13.6
Susquehanna River	10,155	10	178				**47.3**	26	7	1	6.06	2.54	4.0
Susquehanna River	18,299	10	166				**57.7**	20	9	5	4.92	2.55	26.9
Susquehanna River	26,003	10					**61.5**	23	11	3	4.73	2.62	29.9

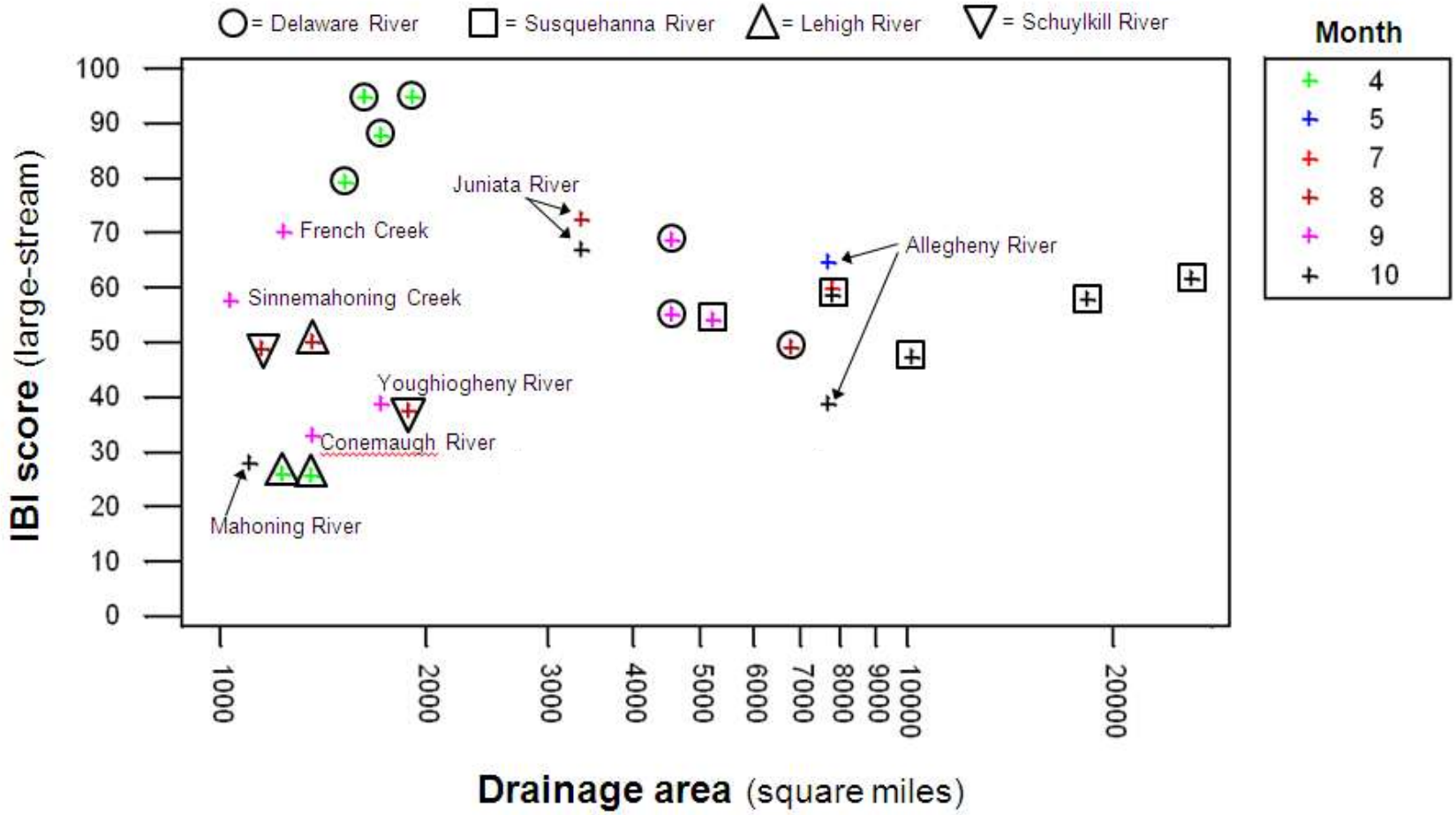


Figure 35. IBI scores for 27 samples from 24 large river sites (i.e. sites with drainage area over 1,000 square miles) coded by sample month. Note logarithmic scale for drainage area.

Dominance

Sometimes, individuals from one taxon or a couple taxa will heavily dominate a sub-sample (i.e., represent more than 33%, 50%, even 67% of all the organisms in the sub-sample). This often occurs in smaller streams in the spring, but can vary depending which taxon or taxa dominate. Frequently, only a handful of taxa heavily dominate sub-samples. Common dominance characteristics for each of these taxa are discussed below.

1. **Chironomidae**

   Chironomidae often dominate springtime samples, but can dominate in just about any season. This dipteran family can dominate samples from streams large and small. Very heavy Chironomidae dominance in wadeable, freestone, riffle-run streams often signals some sort of pollution, commonly organic enrichment and/or sedimentation.

2. **Prosimulium**

   Prosimulium dominance often occurs in March and April during seasonal larval population booms. Early spring dominance by this blackfly genus is often heavy (over 50% of sub-samples) and can occur in relatively pristine streams as well as streams impacted by a variety of human activities, so is not a reliable sign of anthropogenic impact. However, extremely heavy Prosimulium dominance (over 75% of sub-samples) may be a sign of agricultural impacts.

3. **Amphinemura and Leuctra**

   Dominance by either or both of these stonefly genera often occurs in March, April, and May – particularly April and May. Often times, dominance by either or both of these stonefly genera can be heavy (over 50% of sub-samples), which is a fairly reliable sign of acid deposition impacts, especially if observed concurrently with low mayfly abundance and diversity.

4. **Ephemerella**

   Ephemerella dominance often occurs in March, April, and May – particularly April and May. These mayflies can be dominant in larger streams as well as smaller streams. Dominance by this mayfly genus may be a signal of agricultural impacts, but can occur in relatively pristine streams too.

5. **Hydropsychidae** (Diplectrona, Cheumatopsyche, Ceratopsyche, Hydropsyche)

   Dominance by these hydropsychid caddisfly genera more commonly occurs in summer, fall, and early winter than spring. These caddisflies can dominate larger stream samples. Diplectrona dominance is a fairly reliable sign of mining impacts – especially when seen with low mayfly diversity and abundance. Cheumatopsyche, Ceratopsyche, and Hydropsyche dominances are fairly reliable signs of agriculture and/or development impacts.

6. **Stenelmis**

   Dominance by this beetle genus often occurs from late spring through fall and can occur in larger systems as well as smaller systems. Stenelmis dominance is a fairly reliable signal of agricultural impacts, although Stenelmis dominance can occur in more pristine streams that are lower gradient as well.

In addition to the taxa listed above, Baetis, Isonychia, Allocapnia, Oligochaeta, and Gammarus sometimes dominate sub-samples, but much less frequently than those taxa described above.

Because of the sub-sampling procedures used in sample processing, heavy dominance by individuals from one taxon or a few taxa often means that the diversity of organisms in the sub-sample is low compared with the diversity of organisms in the whole sample. Such dominance can drive down individual metric scores – especially Shannon Diversity and metrics based on taxonomic richness – and subsequently the multimetric IBI score. This is of particular concern with Prosimulium because dominance by this blackfly genus can occur in relatively pristine streams, whereas dominance by many of the other commonly-dominant taxa listed above often signals some sort of pollution impact. If a sample is heavily dominated by Prosimulium, it may mean that many taxa present in the sample do not appear in the sub-sample, and the index scores may be unduly low.

Although this phenomenon could be dealt with by altering sub-sampling procedures for heavily-dominated samples, biologists are encouraged to use their best professional discretion when dealing with these situations, and to realize the discussed implications heavy dominance by one taxon or a few taxa may have on the metric and index scores. It may be helpful to document what taxa are present in the entire sample that do not appear in the sub-sample and even to determine rough relative abundances of these taxa in the whole sample to get an idea how much diversity is not represented in the sub-samples. In some instances, additional sampling may be required to confidently assess the stream if an initial sample is heavily dominated by individuals representing one or a few taxa. This especially may be the case with late winter or early spring samples dominated by Prosimulium.

PENNSYLVANIA TIERED AQUATIC LIFE USE WORKSHOPS

Numerous professional aquatic biologists gathered in Harrisburg, Pennsylvania on three separate occasions (August 8 and 9, 2006; August 22 and 23, 2007; May 15 and 16, 2008) to conduct tiered aquatic life use (TALU) workshops. The underlying concepts and procedural details of these workshops are described by Gerritsen and Jessup (2007). The basic idea of the workshop was to assign benthic macroinvertebrate samples to one of a series of biological condition tiers based on experienced biologists voting for different tier assignments. Good agreement among 45 biologists participating in the three TALU workshops and consistency with empirical evidence indicates the conceptual biological condition gradient (BCG) model reflects important aspects of biological condition along a general stressor gradient (Davies and Jackson 2006). Davies and Jackson (2006) promote use of the BCG as a descriptive model of ecosystem response to stress using six conceptual tiers (Figure 36).

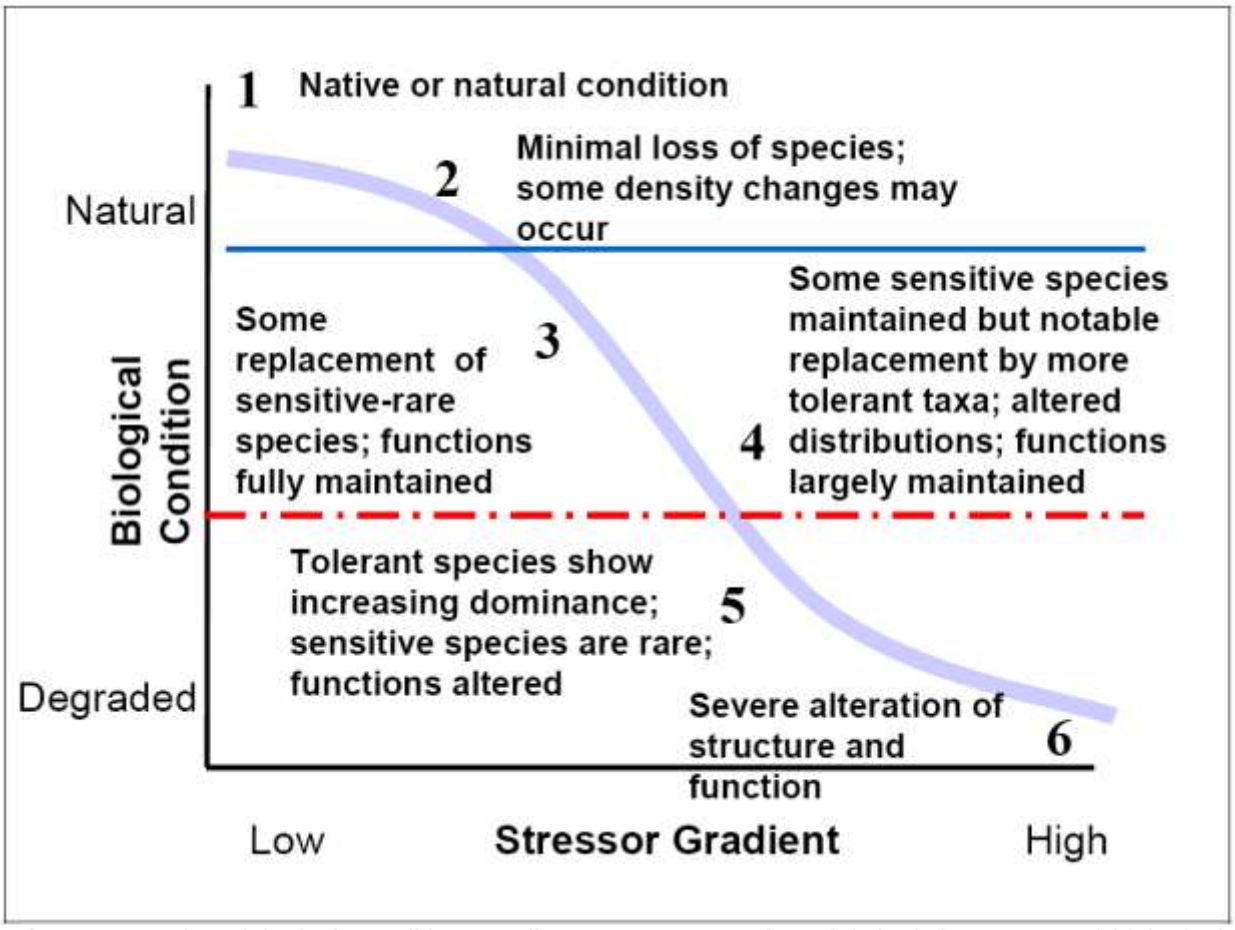


Figure 36. The Biological Condition Gradient – a conceptual model depicting stages of biological condition responses to an increasing stressor gradient – a*dapted from Davies and Jackson (2006)*.

Davies and Jackson (2006) offer that the biological condition required to support an ALU for a specific water body can be described in terms of BCG tiers. For example, the biological condition associated with wild brook trout reproduction requires a very high-quality stream and may be defined as a narrow range of nearly natural BCG tiers, while the biological condition needed to support warm water recreational fisheries may span a broader range of conditions. Davies and Jackson (2006) note that individual applications of the BCG may not require – or be able to distinguish – six tiers, but the BCG development group concluded that six biological condition tiers can be qualitatively distinguished by well-designed and rigorous monitoring programs and that smaller increments of change are useful to show improvements or losses in biological condition.

In addition, many of the biologists who participated in development and testing of the BCG reported that the ecological characteristics conceptually described by tiers 1 through 4 correspond to how they interpret the Clean Water Act interim goal for protection and propagation of aquatic life (Davies and Jackson 2006). Further, the same biologists identified the characteristics described by tiers 1 and 2 as indicative of biological integrity (Davies and Jackson 2006).

Potential pitfalls of the BCG approach include: (1) lack of assessment experience and difficulty of practically and accurately assessing the status of some BCG attributes (e.g., ecosystem function); (2) a consensus definition of tier 1 conditions; and (3) the lack of regionally evaluated species tolerance to general and specific stressors.

The results of the Pennsylvania TALU workshops indicate that professional aquatic biologists from a number of organizations with extensive experience sampling benthic macroinvertebrates and other aquatic life (e.g., fish, periphyton) in the region generally agree on the characteristics exhibited by "reference condition" or "natural" benthic macroinvertebrate communities in the Commonwealth for wadeable, freestone, riffle-run streams. This is an important finding that provides consistent meaning to quantification of these characteristics and decisions based on biological criteria for ALU attainment.

If we apply the large-stream IBI to samples from streams draining more than 50 square miles and the small-stream IBI to samples from streams draining less than 50 square miles, we see very good agreement between IBI scores and mean BCG tier assignments for 92 samples evaluated at the three TALU workshops in Pennsylvania (Figure 37). It should be noted that the IBI scores presented by Gerritsen and Jessup (2007) are based on a different set of metrics than the IBI developed in this report. The IBI scores presented by Gerritsen and Jessup (2007) differ from the IBI presented in the present report in the following ways:

- The standardization value for the Total Taxa Richness metric was 35 in the 2007 IBI.
- The EPT Richness metric in the 2007 IBI was calculated using all EPT taxa rather than only EPT taxa with PTVs of 4 or less. The 2007 standardization value for this metric was 23.
- The 2007 standardization value for the Beck's Index metric was 39.
- The 2007 standardization value for the Shannon Diversity metric was 2.90.
- The 2007 standardization value for the Hilsenhoff Biotic Index metrics was 1.78.
- The Percent Sensitive Individuals metric in the 2007 IBI was calculated using taxa with PTVs from 0 to 5 rather than 0 to 3. The 2007 standardization value for this metric was 92.5.

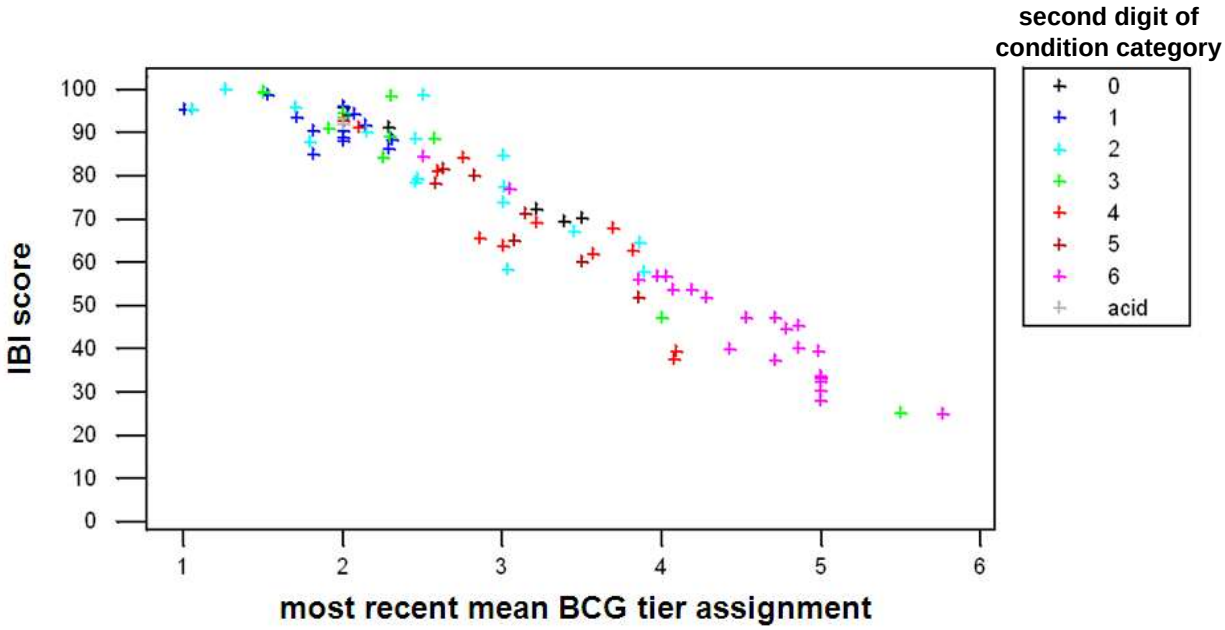


Figure 37. Scatterplot of IBI scores with mean BCG tier assignment from the most recent TALU workshop color-coded by last digit of the condition categories defined in this project. The large-stream IBI was applied to samples from sites draining more than 50 square miles. The small-stream IBI was applied to samples from sites draining less than 50 square miles.

AQUATIC LIFE USE ATTAINMENT BENCHMARKS

For purposes of assessing ALU attainment based on IBI scores, use attainment thresholds or benchmarks can be established for specific stream types, regions and ALU levels. The multimetric index approach offers the ability to use a single index score to simplify management and decision-making (Barbour et al. 1999). The single index value may not determine the exact nature of stressors affecting the ecosystem, but analysis of the individual metrics may offer some insight into causes of ecosystem stress (Barbour et al. 1999). Thus, the index score can be used as a stand-alone assessment tool to represent aquatic life use attainment status, but the assessment process may be strengthened by considering the index score in concert with other available information (Barbour et al. 1999).

The selection of the appropriate criteria heavily depends on the nature of the samples in the dataset, especially the samples used to define the reference condition (Hughes 1995; Barbour et al. 1999; Stoddard et al. 2006). The extremes of biological condition (i.e., severely degraded and nearly pristine conditions) are usually easier to deem acceptable or unacceptable deviations from natural conditions than middle-of-the-road conditions (Hughes 1995). Any set of undisturbed sites will naturally exhibit a range of scores at any point in time (Stoddard et al. 2006), which is why spatial and temporal precision of the index were estimated for this project. Barbour et al. (1999) recommend using established percentiles of multimetric index scores for the reference sites to discriminate between severely degraded and nearly natural conditions. Barbour et al. (1999) also note that the range of index scores can be subdivided into any number of categories corresponding to various levels of degradation or use attainment.

Due to the influences of annual seasons and drainage area seen in the dataset, PADEP recognizes different assessment tools and use attainment thresholds are appropriate for samples collected during different times of the year and from different size stream systems. It is noted that some site-specific exceptions to any thresholds may exist because of local scale natural limitations (e.g., habitat availability) on biological condition (Hughes 1995).

Based on the results of the technical analyses presented above, the results of the TALU workshops, feedback from PADEP biologists and other colleagues, as well as policy considerations, PADEP implements a multi-tiered benchmark decision process for wadeable, freestone, riffle-run streams in Pennsylvania that incorporates stream size and sampling season as factors for determining ALU attainment and impairment based on benthic macroinvertebrate sampling (Figure 38).

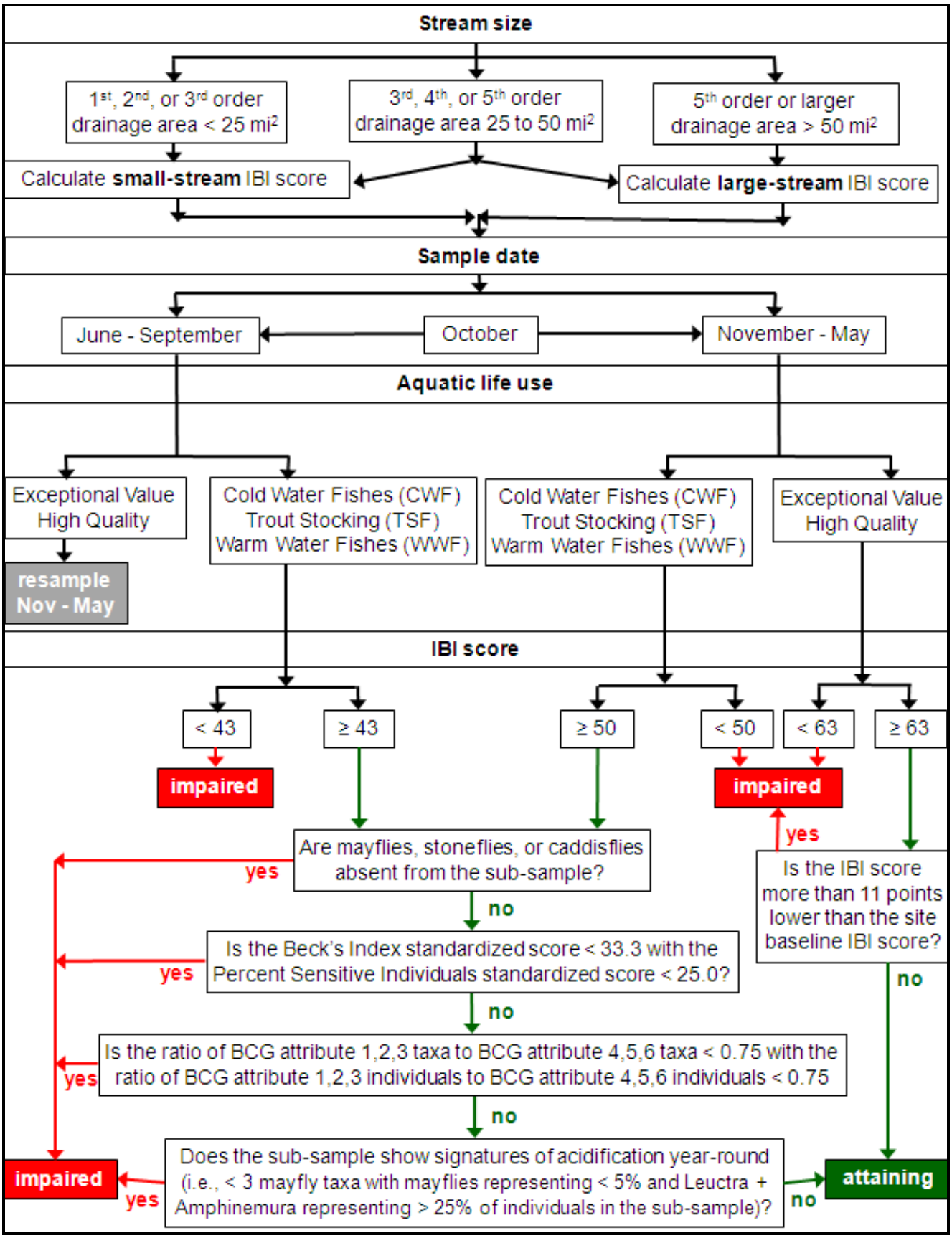

Stream size
1st, 2nd, or 3rd order drainage area < 25 mi²
3rd, 4th, or 5th order drainage area 25 to 50 mi²
5th order or larger drainage area > 50 mi²
Calculate small-stream IBI score
Calculate large-stream IBI score
Sample date
June - September
October
November - May
Aquatic life use
Exceptional Value High Quality
Cold Water Fishes (CWF) Trout Stocking (TSF) Warm Water Fishes (WWF)
Cold Water Fishes (CWF) Trout Stocking (TSF) Warm Water Fishes (WWF)
Exceptional Value High Quality
resample Nov - May
IBI score
< 43
≥ 43
≥ 50
< 50
< 63
≥ 63
impaired
impaired
yes
Are mayflies, stoneflies, or caddisflies absent from the sub-sample?
yes
no
Is the IBI score more than 11 points lower than the site baseline IBI score?
Is the Beck's Index standardized score < 33.3 with the Percent Sensitive Individuals standardized score < 25.0?
yes
no
no
Is the ratio of BCG attribute 1,2,3 taxa to BCG attribute 4,5,6 taxa < 0.75 with the ratio of BCG attribute 1,2,3 individuals to BCG attribute 4,5,6 individuals < 0.75
yes
no
impaired
Does the sub-sample show signatures of acidification year-round (i.e., < 3 mayfly taxa with mayflies representing < 5% and Leuctra + Amphinemura representing > 25% of individuals in the sub-sample)?
yes
no
attaining

Figure 38. A simplified framework for the aquatic life use assessment process. *** Questions 1 and 3 must be applied to small-stream samples collected from November to May, but do not have to be applied to large-stream samples or samples collected from June to September. Although this simplified decision matrix should guide most assessment decisions for benthic macroinvertebrate samples from Pennsylvania's wadeable, freestone, riffle-run streams using the collection and processing methods discussed above, situations exist where this simplified assessment schematic will not apply exactly as outlined – some such situations are discussed in the following text.

The first step in the aquatic life use assessment process for wadeable, freestone, riffle-run streams in Pennsylvania based on benthic macroinvertebrate sampling considers stream size. PADEP does not set a single cutoff drainage area or stream order threshold to define which set of metric standardization values and which resulting IBI (i.e., large-stream or small-stream) should be applied. However – as stated above – data suggest that the small-stream approach is usually appropriate for samples from first, second, and third order streams draining less than 25 square miles of land, while the large-stream approach is usually appropriate for samples from fifth order and larger streams draining more than 50 square miles.

There are many important considerations when deciding whether to apply the small-stream or large-stream metric standardization values to a sample. Many stream systems experience a variety of changes as they flow from headwaters on downstream. These changes include, but are certainly not limited to changes in canopy shading, energy dynamics, algal growth, erosional and depositional patterns, habitat distributions, water temperature, and flow regimes. These shifts manifest themselves uniquely in each watershed. Streams in more northern, high elevation, high relief areas of the state may maintain cooler water, flashier flows, larger-particle substrates, and other characteristics typical of smaller streams at comparable drainage areas or stream orders when compared with streams in more southern, low elevation, low relief areas of the state. Local climatological and geological patterns also affect a stream's character.

When deciding which set of metric standardization values (i.e., small-stream or large-stream) to apply, care should be taken not to conflate human-induced changes to streams with natural landscape and climatological variations. For example, a stream draining 26 square miles of mostly corn and soybean fields with little forested riparian buffer may experience warmer water temperatures and more silted substrates than a stream of similar size draining a more forested watershed. The warmer water and more silted substrates of the agricultural stream may be characteristics typical of larger streams, but if those characteristics are primarily human-induced, then that argues against applying the large-stream metric standardization values based on the presence of those characteristics in the stream.

For streams of intermediate size (i.e., third, fourth, and some fifth order streams draining between 25 and 50 square miles of land), it will often be informative to consider both the small-stream and large-stream IBI scores and associated benchmarks. For example, if a sample from a fourth order site draining 30 square miles scores 77.0 on the small-stream IBI and 90.2 on the large-stream IBI and passes the additional screening questions, both approaches indicate aquatic life use attainment, so the use assessment decision is the same regardless of which set of metric standardization values is applied. In another instance, a sample collected in mid-March from a site draining 36 square miles may score

44.1 on the small-stream IBI – indicating impairment – while scoring 51.2 on the large-stream IBI – indicating possible attainment. Here, the small-stream and large-stream IBI score assessment decisions diverge. In such situations it may be especially useful to consider the additional screening questions – detailed below – when making an assessment decision.

The second step in the aquatic life use assessment process for wadeable, freestone, riffle-run streams in Pennsylvania based on benthic macroinvertebrate sampling considers sampling season. Samples collected during summer and early autumn months (i.e. June through September) are held to different IBI attainment thresholds than samples collected November through May since benthic macroinvertebrate communities in most wadeable, freestone, riffle-run streams in Pennsylvania exhibit consistent patterns of lower taxonomic diversity and organismal abundance during the summer and early autumn months compared with other times of the year. These seasonal index periods are intended as general guidelines and may vary slightly year-to-year depending on local climatological conditions. For example, a sample collected from a low elevation, low latitude stream during the last week of May in a particularly hot, dry year may be more properly evaluated using procedures set forth for the summer months – especially if many mayflies have already emerged from the stream – while a sample collected from a high elevation, high latitude location during the first week of June in a uncharacteristically cool, wet year may be more properly evaluated using the November to May procedures – especially if many mayfly nymphs are still present in the benthos.

October often is a transitional time for benthic macroinvertebrate communities in Pennsylvania with samples from earlier in the month resembling late summer communities (e.g., relatively low diversity and abundance) and samples from later in the month resembling early winter communities (e.g., increasing abundance of winter stoneflies). Therefore, depending on local climate, basin geology, and other factors discussed above (e.g., latitude, elevation, basin relief) samples from October may be evaluated using the June to September benchmarks or the November to May benchmarks. PADEP advises against sampling in mid-October to avoid these issues. In fact, *PADEP encourages sampling be conducted in the November to May time frame whenever possible.*

For samples collected between November and May, IBI scores < 50 result in aquatic life use impairment. Samples collected during these months scoring ≥ 50 on the appropriate IBI are subject to four screening questions before the aquatic life use can be considered attaining. **These additional screening questions are:**

1. **Are mayflies, stoneflies, or caddisflies absent from the sub-sample?** Organisms representing these three taxonomic orders are usually found in most healthy wadeable, freestone, riffle-run streams in Pennsylvania. If any or all of these orders are absent from a sample, this strongly suggests some sort of anthropogenic impact. Samples where one of these taxonomic orders is absent due to natural conditions (e.g., mayflies absent from a low-pH tannic stream) should be evaluated accordingly. *This question must be applied to small-stream samples collected between November and May, but does not have to be applied to samples from larger streams and samples collected between June and September.*

2. **Is the standardized metric score for the Beck's Index metric < 33.3 with the standardized metric score for the Percent Sensitive Individuals metric < 25.0?**

Although these two metrics go into the IBI calculations, this screening question serves to double check that a sample has substantial richness and abundance of the most sensitive organisms. This question arose from observing that the Beck's Index metric is less sensitive at the lower end of its range and the Percent Sensitive Individuals metric is less sensitive at the upper end of its range. When both these metrics score relatively low, it serves as strong confirmation of impairment. *This question must be applied to all samples.*

3. **Is the ratio of BCG attribute 1,2,3 taxa to BCG attribute 4,5,6 taxa < 0.75 with the ratio of BCG attribute 1,2,3 individuals to BCG attribute 4,5,6 individuals < 0.75?** This screening question evaluates the balance of pollution tolerant organisms with more sensitive organisms in terms of taxonomic richness and organismal abundance. By using the BCG attributes to measure pollution tolerance, this screening question serves as a check against the IBI metrics which account for pollution sensitivity based only on PTVs. *This question must be applied to small-stream samples collected between November and May, but can be relaxed for samples from larger streams and samples collected between June and September.*

4. **Does the sub-sample show signatures of acidification year-round?** The primary acidification signatures in a sub-sample include low mayfly abundance and low mayfly diversity (i.e., scarce mayfly individuals and few mayfly taxa), especially when combined with high abundance of Amphinemura and/or Leuctra stoneflies, occasionally combined with high abundance of Simuliidae and/or Chironomidae individuals. A sub-sample with < 3 mayfly taxa, < 5% mayfly individuals, and > 25% Leuctra and/or Amphinemura stoneflies indicates likely acidification impacts. Acidification effects on benthic macroinvertebrate communities are often most pronounced in small streams with low buffering capacity during the spring months when snowpacks melt and vernal rains are frequent. While it can be difficult to determine if low pH conditions in a stream are natural or more attributable to anthropogenic acidification, sampling of water chemistry and/or fish communities (see Appendix F of PADEP 2009b) in addition to benthic macroinvertebrate communities can help inform assessment of acidic in-stream conditions. With this protocol, PADEP will only impair sites that show persistent acidification signatures year-round. In other words, if a sample has no mayflies and is dominated by Leuctra and Amphinemura in the spring, but a November sample from the same site contains three or more mayfly taxa or over five percent mayfly individuals, the aquatic life use will not be considered impaired because the stream exhibits the ability to recover biological integrity in the fall and winter months. If a spring sample shows acidification signatures, a late fall or early winter sample must be collected before making an aquatic life use assessment decision. *This question must be applied to all samples.*

If the answer to any of the required screening questions is yes for a sample collected between November and May with an IBI score ≥ 50, then the sample is considered impaired without compelling reasons otherwise. If the answer to all of these questions is no for a sample collected between November and May with an IBI score ≥ 50, then the aquatic life use represented by the sample can be considered attaining unless other information (e.g., water chemistry) indicates the aquatic life use may not be fully supported at that location.

For samples collected between June and September, the same logic applies as for samples collected between November and May, but the attainment/impairment threshold is lowered to 43 instead of 50. The 43 benchmark was selected based on analysis of seasonal IBI fluctuations at a number of sites. These analyses showed that sites with relatively high IBI scores (i.e., above 50) during the November to May time frame very rarely had IBI scores drop below 43 during the summer and early autumn months (Figure 39 – see the Big

Wapwallopen Creek sample in particular; many similar plots were evaluated to establish the 43 June to September benchmark). Thus, a June to September benchmark of 43 should prevent assessment decisions to impair a stream's ALU based on summer or early autumn samples when samples from other times of the year indicate the stream is supporting its ALU, while still maintaining the ability to detect ALU impairments that persist year-round.

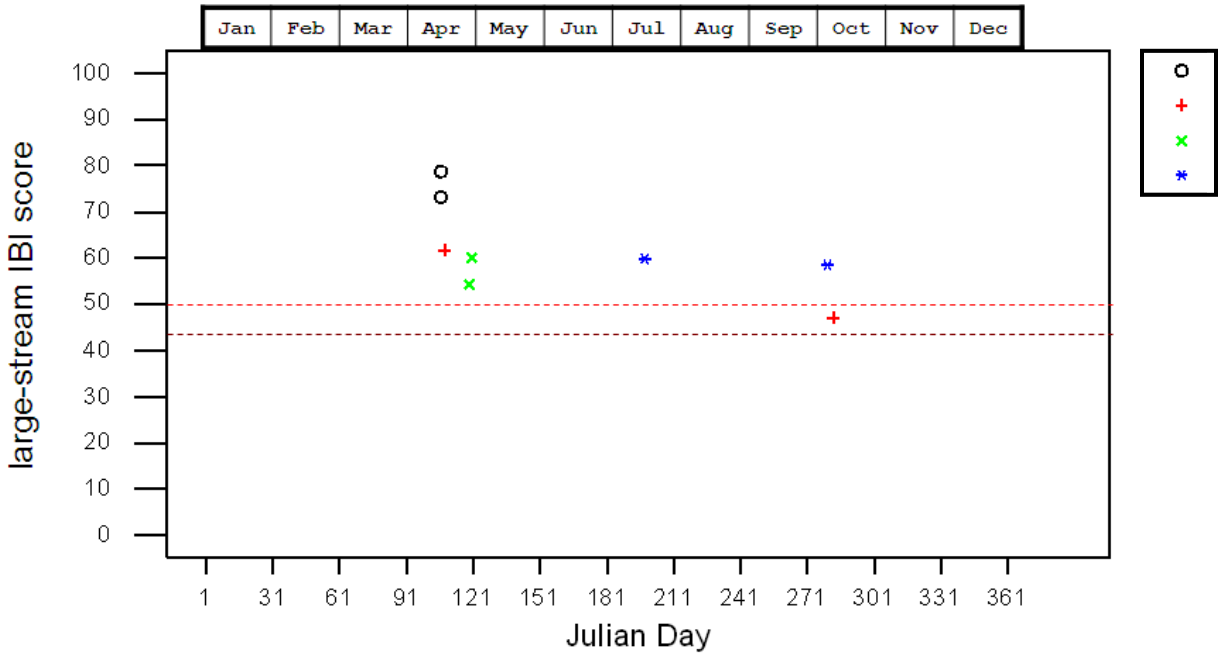
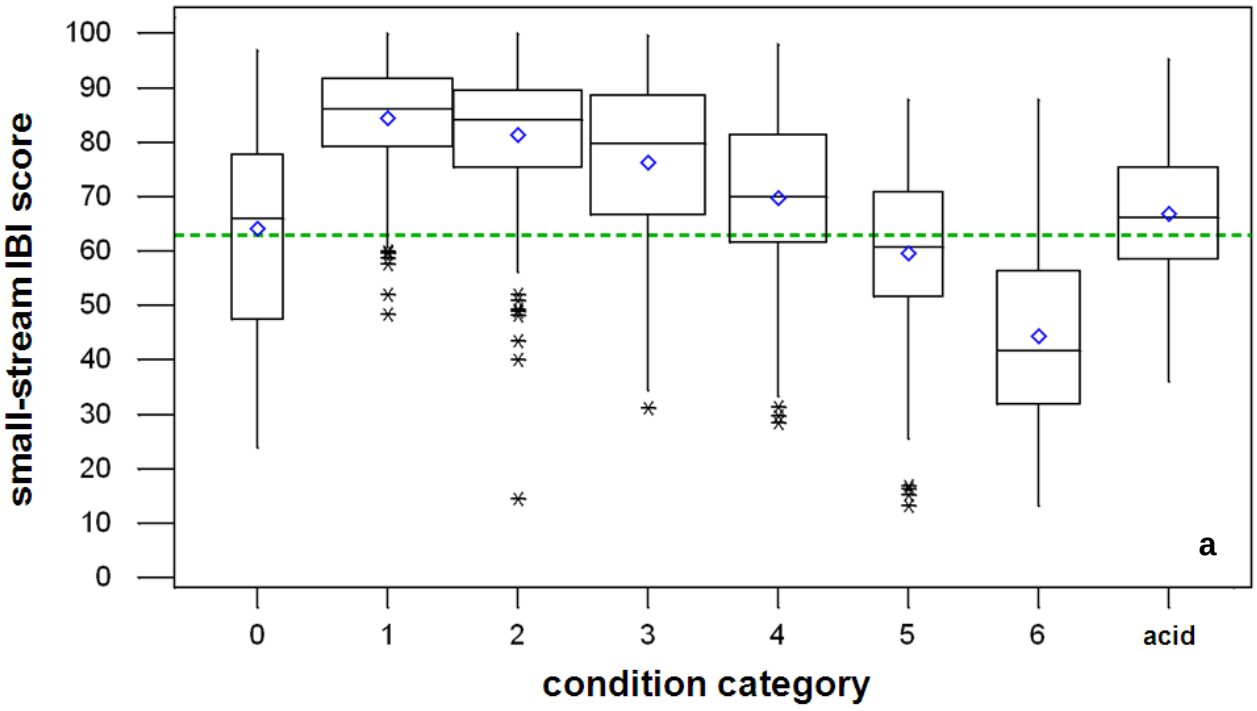


Figure 39. Large-stream IBI scores plotted by Julian Day of sample collection for four sites. Lines drawn at the 50 and 43 benchmarks for visual emphasis. (Black circle = Tunkhannock Creek at 188 square miles. Red plus = Big Wapwallopen Creek at 53 square miles. Green X = Towanda Creek at 66 square miles. Blue asterisk = Susquehanna River at 7,792 square miles.)

For samples collected in the summer and early autumn time frame, the absence of mayflies – and in some instances stoneflies – in samples collected immediately after seasonal hatches may be relaxed. Because benthic diversity may be underrepresented in summer and early autumn samples PADEP encourages monitoring in the November to May time frame if possible. Benthic macroinvertebrate sampling for determining aquatic life use support should only be conducted from June to early October if sampling during other seasons is not possible due to hazardous conditions such as high, fast stream flow.

By combining the ALU-specific IBI benchmarks with the additional ALU assessment screening questions, the ALU assessment decision process outlined above provides for protection of the least impacted wadeable, freestone, riffle-run streams in Pennsylvania at a high level of biotic integrity, while recognizing impacted streams as having impaired ALUs (Figures 40-42, Table 27). These ALU assessment procedures are applied specific to each of Pennsylvania's five ALUs and calibrated according to stream size and sampling season.


small-stream IBI score
100
90
80
70
60
50
40
30
20
10
0
a
0
1
2
3
4
5
6
acid
condition category

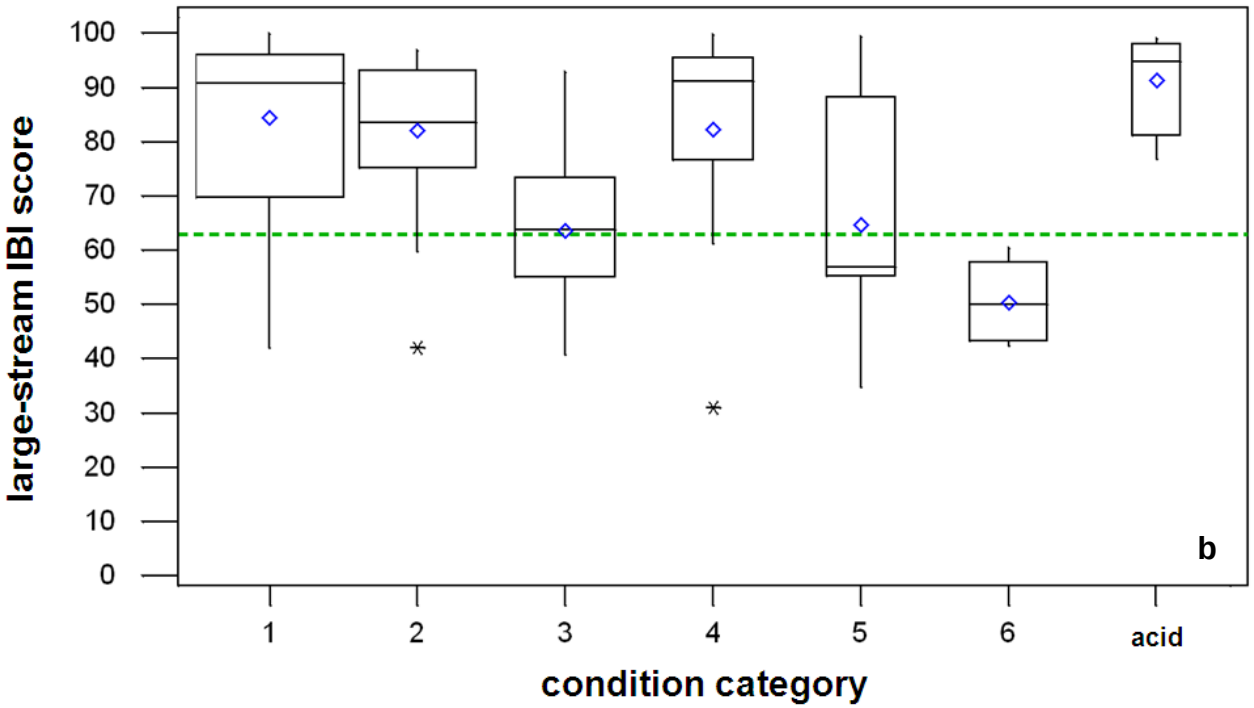
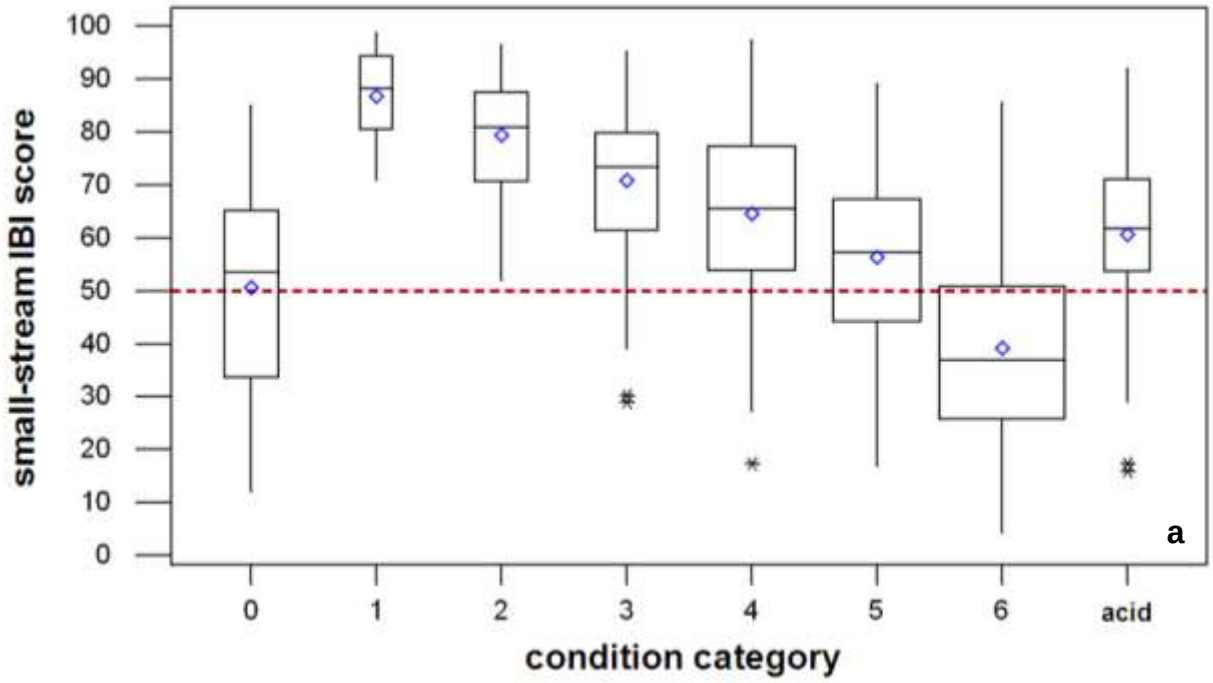


Figure 40. Distribution of (a) small-stream and (b) large-stream IBI scores by condition category for samples from EV and HQ streams in the November to May time frame. Total number of samples in each plot are (a) n = 1,186 and (b) n = 110. For simplicity, the small-stream IBI was applied to samples from sites draining less than 50 square miles and the large-stream IBI was applied to samples from sites draining more than 50 square miles. The antidegradation ALU benchmark of 63 is drawn in for visual reference.



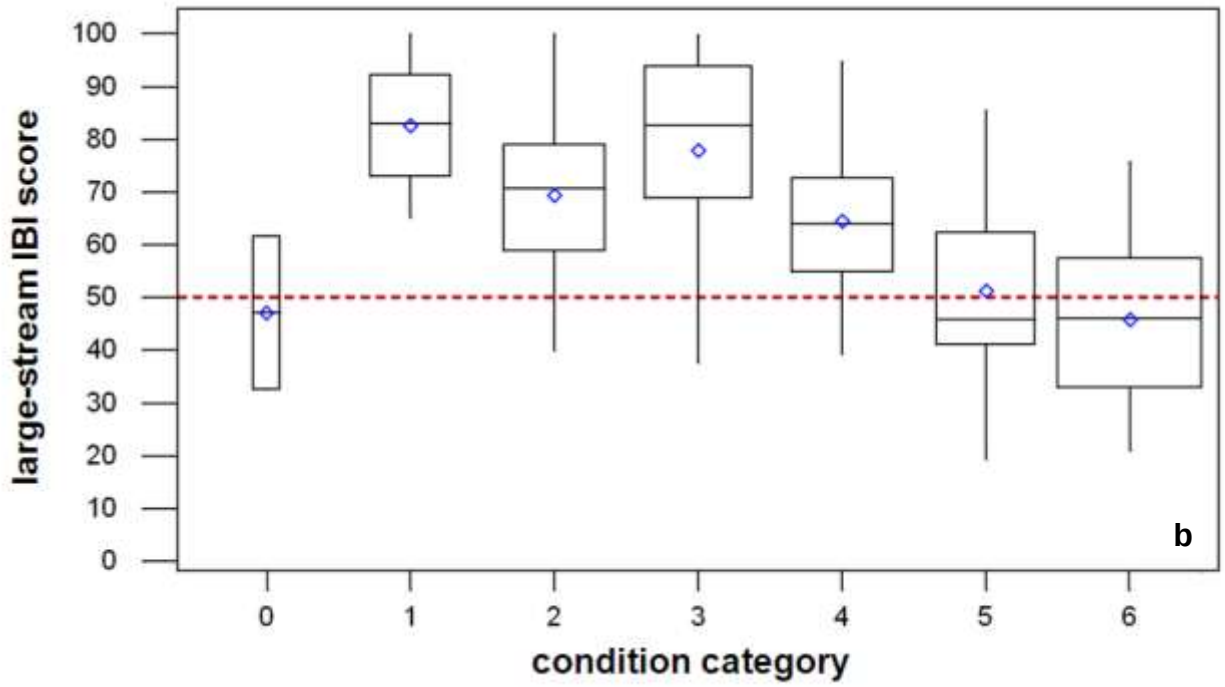
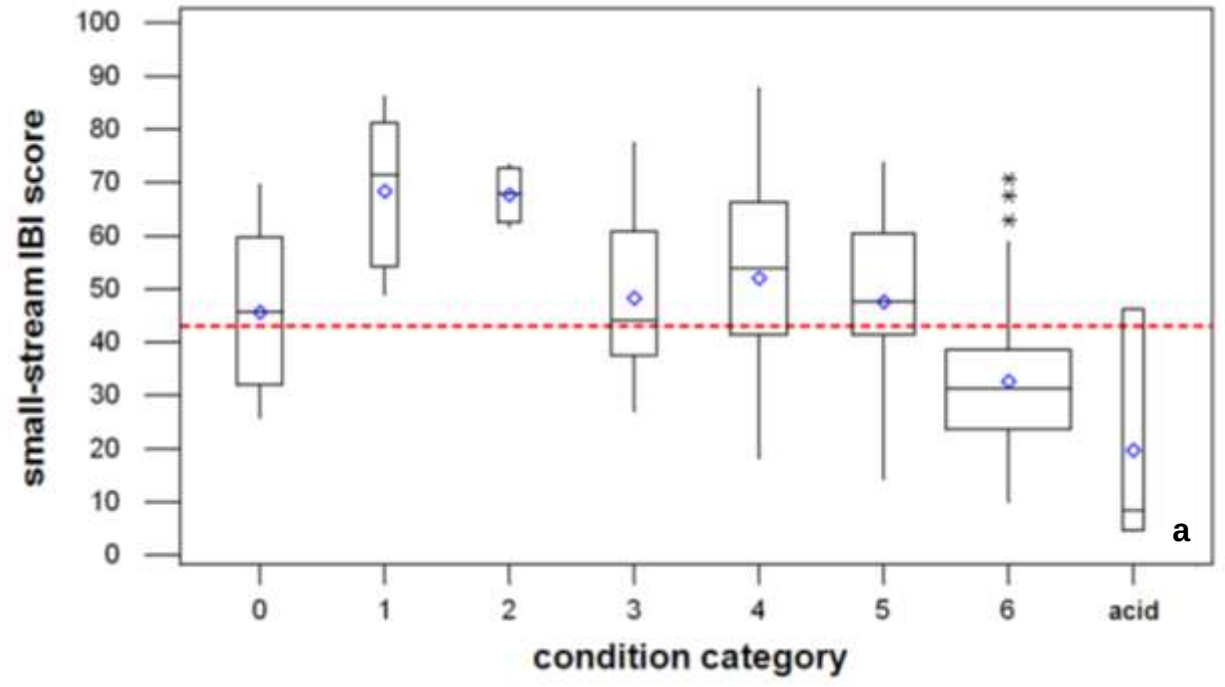


Figure 41. Distribution of (a) small-stream and (b) large-stream IBI scores by condition category for samples from CWF, TSF, and WWF streams in the November to May time frame. Total number of samples in each plot are (a) n = 924 and (b) n = 195. For simplicity, the small-stream IBI was applied to samples from sites draining less than 50 square miles and the large-stream IBI was applied to samples from sites draining more than 50 square miles. The ALU benchmark of 50 is drawn in for visual reference.



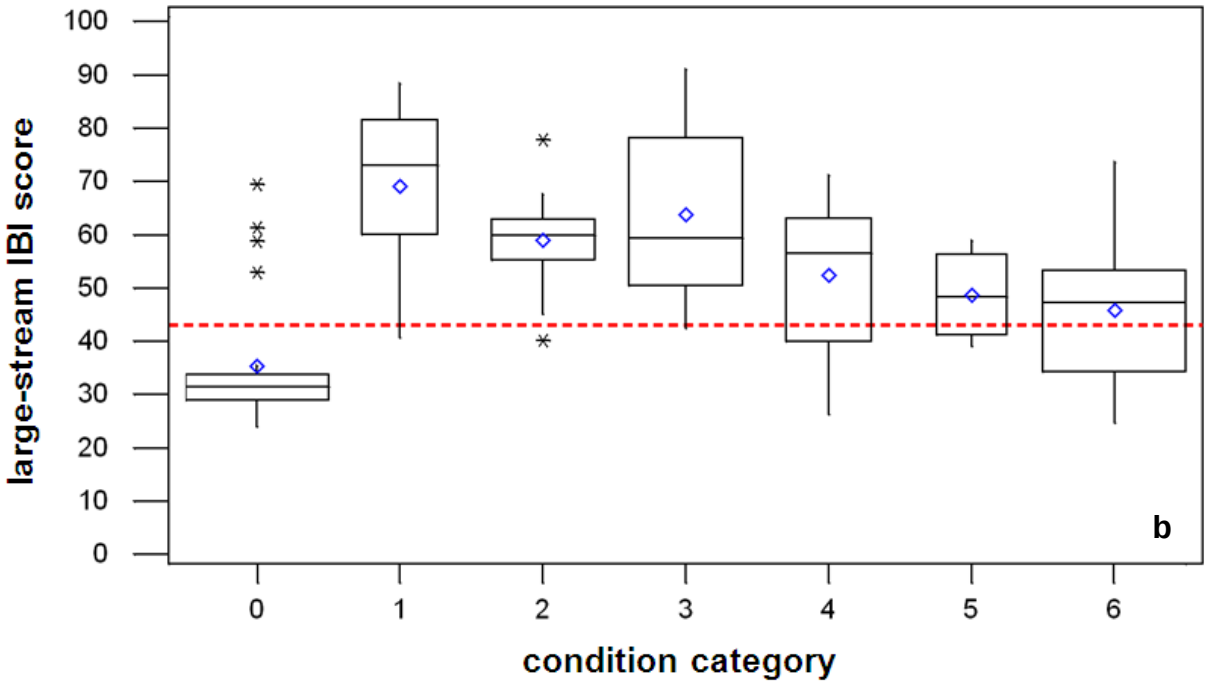


Figure 42. Distribution of (a) small-stream and (b) large-stream IBI scores by condition category for samples from CWF, TSF, and WWF streams in the June to September time frame. Total number of samples in each plot are (a) n = 214 and (b) n = 101. For simplicity, the small-stream IBI was applied to samples from sites draining less than 50 square miles and the large-stream IBI was applied to samples from sites draining more than 50 square miles. The ALU benchmarks of 43 is drawn in for visual reference.

Table 27. Assessment decision results by condition category, stream size, and sampling season. For simplicity, the small-stream procedures were applied to samples from sites draining less than 50 square miles and the large-stream procedures were applied to samples from sites draining more than 50 square miles.

Stream Size	Season	Uses	Assessment Decision	condition category							
				0	1	2	3	4	5	6	acid
small	November to May	HQ EV	# Attaining	22	240	220	158	97	31	13	90
			# Impaired	17	11	20	38	42	41	89	57
			# Total	39	251	240	196	139	72	b102	147
			% Attaining	**56%**	**96%**	**92%**	**81%**	**70%**	**43%**	**13%**	**61%**
			% Impaired	**44%**	**4%**	**8%**	**19%**	**30%**	**57%**	**87%**	**39%**
		CWF TSF WWF	# Attaining	19	21	57	66	96	74	45	11
			# Impaired	43	0	1	19	65	87	287	33
			# Total	62	21	58	85	161	161	332	44
			% Attaining	**31%**	**100%**	**98%**	**78%**	**60%**	**46%**	**14%**	**25%**
			% Impaired	**69%**	**0%**	**2%**	**22%**	**40%**	**54%**	**86%**	**75%**
	June to September	CWF TSF WWF	# Attaining	8	5	4	8	13	14	13	0
			# Impaired	7	0	0	8	11	16	104	3
			# Total	15	5	4	16	24	30	117	3

			% Attaining	53%	100%	100%	50%	54%	47%	11%	0%
			% Impaired	47%	0%	0%	50%	46%	53%	89%	100%
large	November to May	HQ EV	# Attaining	0	33	21	9	9	2	0	4
			# Impaired	0	4	2	8	2	6	10	0
			# Total	0	37	23	17	11	8	10	4
			% Attaining	0%	89%	91%	53%	82%	25%	0%	100%
			% Impaired	0%	11%	9%	47%	18%	75%	100%	0%
		CWF TSF WWF	# Attaining	1	18	26	28	24	11	16	0
			# Impaired	1	0	3	4	5	16	42	0
			# Total	2	18	29	32	29	27	58	0
			% Attaining	50%	100%	90%	88%	83%	41%	28%	0%
			% Impaired	50%	0%	10%	13%	17%	59%	72%	0%
	June to September	CWF TSF WWF	# Attaining	3	6	10	12	6	2	6	0
			# Impaired	22	1	3	4	3	4	19	0
			# Total	25	7	13	16	9	6	25	0
			% Attaining	12%	86%	77%	75%	67%	33%	24%	0%
			% Impaired	88%	14%	23%	25%	33%	67%	76%	0%
overall			# Attaining	53	323	338	281	245	134	93	105
			# Impaired	90	16	29	81	128	170	551	93
			# Total	143	339	367	362	373	304	644	198
			% Attaining	37%	95%	92%	78%	66%	44%	14%	53%
			% Impaired	63%	5%	8%	22%	34%	56%	86%	47%

Limestone Influence

As discussed in the introduction, PADEP deploys a different sampling methodology and assessment protocol for limestone spring streams whose flow is mostly or entirely derived from groundwater in areas with substantial primary calcareous geologies (PADEP 2009a) than for freestone streams. The sampling methodology and assessment protocol for these limestone spring streams incorporate the understanding that streams in areas receiving a substantial amount of flow from groundwater attributable to karst geologies often naturally have less diverse benthic macroinvertebrate communities than streams draining freestone geologies. This lower benthic macroinvertebrate community diversity in limestone spring streams is attributable in large part to less variable flow and thermal characteristics of such systems when compared with freestone streams that often exhibit flashier flows and a wider range of temperatures.

Some streams in Pennsylvania drain basins underlain partially by freestone geologies and partially by calcareous geologies. Such streams are often encountered in central regions of the state – especially in upper portions of the Juniata River basin – where they drain sandstone and/or quartzite upland ridges, fairly steep shale slopes, and lower gradient calcareous valley floors. The calcareous valley geologies in these basins contributes to relatively high alkalinities and relatively high and consistent base flows in streams – characteristics of limestone spring streams – when compared with streams draining basins

with no calcareous geologies. However, the upland sandstone, quartzite, and shale areas of these basins often contribute substantial surface runoff, which leads to surges in flow during rainfall and snowmelt events and dilution of alkalinity derived from the calcareous valleys. These streams – often referred to as "limestone-influenced" – exhibit some characteristics of limestone spring streams and some characteristics of freestone streams.

We often see substantial agriculture in the fertile valleys of these limestone-influenced streams, which makes it difficult to definitively establish reference conditions specific to these unique streams. However, there is evidence that the benthic macroinvertebrate communities in limestone-influenced streams are naturally less diverse than in freestone streams of similar size and with similar land uses. This lower diversity of benthic macroinvertebrate communities in limestone-influenced streams likely reflects the less variable flow and thermal patterns in these streams caused by the stabilizing influence of the substantial groundwater flowing into the streams through the calcareous valley geologies. Commonly, the benthic macroinvertebrate communities in limestone-influenced streams exhibit relatively low stonefly diversity and abundance when compared with streams of similar size and condition that drain freestone geologies.

In light of these considerations, use attainment benchmarks may be justifiably relaxed for samples from limestone-influenced streams. The June to September IBI benchmark of 43 for freestone streams can be applied to limestone-influenced streams year-round, but the four screening question should still be applied as outlined above to samples from limestone-influenced streams to make ALU assessment decisions.

Antidegradation, Special Protection Considerations

The assessment decision process is somewhat different for streams with special protection uses of high-quality (HQ) or exceptional value (EV) waters. PADEP will protect special protection streams based on a baseline IBI score determined by previous surveys. Subsequent samples from HQ and EV streams will be compared to the baseline IBI score for a given site using the IBI temporal precision estimates (Table 23). For example, if Riverkill Creek is designated HQ and a previous sample from a given site on Riverkill Creek using the protocol described above results in a mid-April IBI score of 78.0, this IBI score of 78.0 would be the baseline IBI score for that site. Future samples from that site collected November to May that score more than 10.0 IBI points below 78.0, would be considered impaired. Since PADEP's sampling season for special protection surveys is November to May, we need not be concerned about how June to October samples compare to the baseline IBI – PADEP will only make assessment decisions for HQ and EV streams based on samples collected November to May. The temporal precision estimate of 10.0 points is used because it approximates the October to May temporal precision estimate calculated above (Table 23). PADEP will apply the more restrictive March to May and October to February temporal precision estimates – about 9.0 and 8.0 IBI points, respectively – to special protection use assessments if the situation is appropriate (e.g., if the baseline IBI was established in April, future March to May samples that score more than 9.0 points lower than the baseline will be considered impaired). Furthermore, any sample from an HQ or EV stream that scores less than 63.0 on the IBI will be considered impaired without compelling reasons otherwise (e.g., a stream was designated HQ or EV for a reason other than assessment of the benthic macroinvertebrate community).

Applications and Exceptions

If a sample results in fewer than 160 total organisms in the entire sample, the IBI and assessment procedures may not apply exactly as outlined above. The IBI and associated benchmarks are calibrated for use with sub-samples containing 160 to 240 organisms, so applications of the IBI to samples containing less – or more – than the target number of organisms, cannot necessarily be assessed using the procedures and benchmarks outlined above. Low abundance of benthic organisms often indicates toxic pollution or severe habitat alterations, which must be considered in making holistic stream assessments.

The use assessment decision processes set forth above are intended as general guidelines, not as hard-and-fast rules. The procedures and guidelines discussed above will provide tenable assessments – as required by federal and state law – of benthic macroinvertebrate community conditions for the vast majority of samples collected from wadeable, freestone, riffle-run streams in Pennsylvania. However, as noted by Hughes (1995), there will be exceptional circumstances – such as those outlined in the Pennsylvania Code (2011: Title 25, Section 93.4.(b) relating to less restrictive uses) – when the above assessment procedures do not apply (e.g., there are no obvious sources of impairment and natural factors such as habitat availability or water chemistry limit biotic potential). In some situations a biologist's local knowledge of conditions may warrant a decision not arrived at using these guidelines. As discussed above, the use assessment procedures outlined in this report should be applied with care to samples from large rivers (i.e., rivers draining more than 1,000 square miles of land) because of the limited dataset of samples available on such rivers. In other situations, like when samples are heavily dominated by Prosimulium larvae – as discussed above – often times this will unduly lower metric and IBI scores, confounding the assessment decision procedures outlined above. In such situations, the investigating biologist may have to re-sample the site after the seasonal Prosimulium larval boom, or the biologist may have to rely on a more qualitative analysis of metric scores, sample composition, and site conditions to arrive at an assessment decision. In any instance, evaluating stream samples requires mindfulness of particular conditions, and is not always a definite, exact exercise. A certain section of stream may represent a transition between pool-glide, low-relief, marshy, glaciated uplands where the substrate is mostly fine-grained sand and higher-gradient lower reaches filled with cobble-strewn riffles and runs. Some years see cooler, wetter springs than other years. Nevertheless, for the vast majority of cases involving benthic macroinvertebrate samples from wadeable, freestone (and limestone-influenced), riffle-run streams in Pennsylvania using the protocols described above, the assessment procedures described in this report will lead to tenable ALU assessment decisions.

REFERENCES

Anderson, M.A. and A.P. Olivero. 2003. The Nature Conservancy stream macrohabitats: lower New England ecoregional plan. The Nature Conservancy.

Barbour, M.T., J. Gerritsen, B.D. Snyder, and J.B. Stribling. 1999. Rapid bioassessment protocols for use in streams and wadeable rivers: periphyton, benthic macroinvertebrates and fish, second edition. EPA 841-B-99-002. United States Environmental Protection Agency; Office of Water; Washington, D.C.

Barbour, M.T., J.B. Stribling, and J.R. Karr. 1995. Multimetric approach for establishing biocriteria and measuring biological condition. Chapter 6 in *Biological assessment and criteria: tools for water resource planning and decision making*, W.S. Davis and T.P. Simon, eds. (pp. 63 – 77). CRC Press, Boca Raton.

Beck, W.H., Jr. 1955. Suggested method for reporting biotic data. Sewage and Industrial Waste 27(10): 1193-1197.

Burton, J. and J. Gerritsen. 2003. A stream condition index for Virginia non-coastal streams. Prepared for United States Environmental Protection Agency, Office of Water, Region 3 Environmental Services Division, and Virginia Department of Environmental Quality. Tetra Tech, Inc. Owings Mills, Maryland.

Carlisle, D.M., M.R. Meador, S.R. Moulton II, and P.M. Ruhl. 2007. Estimation and application of indicator values for common macroinvertebrate genera and families of the United States. Ecological Indicators 7(2007): 22-33.

Ciolkosz, E.J. and E.R. Levine. 1983. Evaluation of acid rain sensitivity of Pennsylvania soils. Research Project Technical Completion Report A-058-PA. Institute for Research on Land and Water Resources, The Pennsylvania State University, University Park, Pennsylvania.

Davies, S.P. and S.K. Jackson. 2006. The biological condition gradient: a descriptive model for interpreting change in aquatic ecosystems. Ecological Applications 16(4):1251-1266.

Davis, W.S. and T.P. Simon. 1995. Introduction to *Biological assessment and criteria: tools for water resource planning and decision making*, W.S. Davis and T.P. Simon, eds. (pp. 3 – 6). CRC Press, Boca Raton.

Fisher, R.A. 1936. The use of multiple measurements in taxonomic problems. Annals of Eugenics 7: 179-188.

Flotemersch, J.E., J.B. Stribling, and M.J. Paul. 2006. Concepts and approaches for the bioassessment of non-wadeable streams and rivers. EPA/600/R-06/127. U.S. Environmental Protection Agency, Office of Research and Development. Washington, D.C.

Gawler, S.C., M.G. Anderson, A.P. Olivero, and M. Clark. 2008. The Northeast Habitat Classification and Mapping Project Final Report. Boston, Massachusetts.

Gerritsen, J., J. Burton, and M.T. Barbour. 2000. A stream condition index for West Virginia wadeable streams. Prepared for United States Environmental Protection Agency, Office of Water, Region 3 Environmental Services Division, and West Virginia Department of Environmental Protection. Tetra Tech, Inc. Owings Mills, Maryland.

Gerritsen, J. and B. Jessup. 2007. Identification of the biological condition gradient for freestone (non-calcareous) streams of Pennsylvania. Prepared for United States Environmental Protection Agency and Pennsylvania Department of Environmental Protection. Tetra Tech, Inc. Owings Mills, Maryland. 53 pp.

Giberson, D.J. and R.J. Mackay. 1991. Life history and distribution of mayflies (Ephemeroptera) in some acid streams in south central Ontario, Canada. Canadian Journal of Zoology 69: 899-910.

Griffith, M.B., S.A. Perry, and W.B. Perry. 1995. Macroinvertebrate communities in headwater streams affected by acidic precipitation in the central Appalachians. Journal of Environmental Quality 24(2): 233-238.

Griffith, M.B., S.A. Perry, and W.B. Perry. 1994. Secondary production of macroinvertebrate shredders in headwater streams with different baseflow alkalinity. Journal of the North American Benthological Society 13(3): 345-356.

Guerold, F., J. Boudot, G. Jacquemin, D. Vein, D. Merlet, and J. Rouiller. 2000. Macroinvertebrate community loss as a result of headwater stream acidification in the Vosges Mountains. Biodiversity and Conservation 9: 767-783.

Hand, D.J. 1981. *Discrimination and classification*. John Wiley and Sons, New York.

Hawkins, C.P. and R.H. Norris. 2000. Performance of different landscape classifications for aquatic bioassessments: introduction to the series. Journal of the North American Benthological Society 19(3): 367-369.

Hawkins, C.P., R.H. Norris, J.N. Hogue, and J.W. Feminella. 2000. Development and evaluation of predictive models for measuring the biological integrity of streams. Ecological Application 10(5): 1456-1477.

Hawkins, C.P. 2006. Quantifying biological integrity by taxonomic completeness: its utility in regional and global assessments. Ecological Applications 16(4): 1277-1294.

Hilsenhoff, W.L. 1977. Use of arthropods to evaluate water quality of streams. Technical Bulletin Number 100. Wisconsin Department of Natural Resources. 15 pp. Madison, Wisconsin.

Hilsenhoff, W.L. 1987. An improved biotic index of organic stream pollution. The Great Lakes Entomologist 20(1): 31-39.

Hilsenhoff, W.L. 1988. Rapid field assessment of organic pollution with a family-level biotic index. Journal of the North American Benthological Society 7(1): 65-68.

Homer, C., C. Huang, L. Yang, B. Wylie, and M. Coan. 2004. Development of a 2001 National Landcover Database for the United States. Photogrammetric Engineering and Remote Sensing 70(7): 829-840.

Hughes, R.M. 1995. Defining acceptable biological status by comparing with reference conditions. Chapter 4 in *Biological assessment and criteria: tools for water resource planning and decision making*, W.S. Davis and T.P. Simon, eds. (pp. 31 – 47). CRC Press, Boca Raton.

Karr, J.R. and D.R. Dudley. 1981. Ecological perspective on water quality goals. Environmental Management 5: 55-68.

Klemm, D.J, P.A. Lewis, F. Fulk, and J.M. Lazorchak. 1990. *Macroinvertebrate field and laboratory methods for evaluating the biological integrity of surface waters*. Environmental Monitoring Systems Laboratory, United States Environmental Protection Agency. Cincinnati, Ohio. EPA-600-4-90-030.

Kimmel, W.G. 1999. Macroinvertebrate community responses to episodic stream acidification on the Laurel Hill in southwestern Pennsylvania. Chapter 4 in The effects of acidic deposition on aquatic ecosystems in Pennsylvania, eds. W.E. Sharpe and J.R. Drohan. Proceedings of the 1998 Pennsylvania

acidic deposition conference, Volume II. Environmental Resources Research Institute, The Pennsylvania State University, University Park, Pennsylvania.

Kruskal, J.B. and M. Wish. 1978. *Multidimensional Scaling*. Sage University paper series on Quantitative Applications in the Social Sciences 07 – 011. Sage Publications. Beverly Hills and London.

Lance, G.N. and W.T. Williams. 1967. A general theory of classificatory sorting strategies – 1. Hierarchical systems. Computer Journal 9: 373-380.

Lenat, D.R. and D.L. Penrose. 1996. History of the EPT taxa richness metric. Bulletin of the North American Benthological Society 13(2).

Ludwig, J.A., and J.F. Reynolds. 1988. *Statistical ecology*. John Wiley and Sons. New York, New York.

Lynch, J.A., K.S. Horner, J.W. Grimm, E. Boyer, and H. Carrick. 2007. Reductions in Acidic Wet Deposition in Pennsylvania Following Implementation of the Clean Air Act Amendments of 1990: 1995-2006. Penn State Institutes of Energy and the Environment – The Pennsylvania State University. University Park, Pennsylvania.

Madarish, D.M. and W.G. Kimmel. 2000. Benthic macroinvertebrate community structure in relation to seasonal and geochemical changes in a chronically acidified stream. Journal of Freshwater Ecology 15(1): 13-27.

Marchant, R. 1999. How important are rare species in aquatic community ecology and bioassessment? A comment on the conclusions of Cao et al. 1999. Limnology and Oceanography 44: 1840-1841.

Marchant, R. 2002. Do rare species have any place in multivariate analysis for bioassessment? Journal of the North American Benthological Society 21(2): 311-313.

Mattson, K.M. and P.L. Angermeier. 2007. Integrating human impacts and ecological integrity into a risk-based protocol for conservation planning. Environmental Management 39(1): 125-138.

Milligan, G.W. 1989. A study of the beta-flexible clustering method. Multivariate Behavioral Research 24: 163-176.

Novotny, V. 2004. Linking pollution to water body integrity – literature review. Center for Urban Environmental Studies, Northeastern University. Boston, Massachusetts.

Pennsylvania Code. Accessed June 24, 2011. http://www.pacode.com/.

Pennsylvania Department of Environmental Protection. 2007. Multihabitat stream assessment protocol. Harrisburg, Pennsylvania.

Pennsylvania Department of Environmental Protection. 2009a. An index of biological integrity for “true” limestone streams. Harrisburg, Pennsylvania.

Pennsylvania Department of Environmental Protection. 2009b. Survey Protocol – Instream Comprehensive Evaluation Surveys. 391-3200-001. Harrisburg, PA.

Pond, G.J. 2010. Patterns of Ephemeroptera taxa loss in Appalachian headwater streams (Kentucky, USA). Hydrobiologia 641: 184-201.

Pond, G.J., S.M. Call, J.F. Brumley, and M.C. Compton. 2003. The Kentucky macroinvertebrate bioassessment index: derivation of regional narrative ratings for assessing wadeable and headwater streams. Kentucky Department for Environmental Protection, Water Quality Branch, Ecological Support Section. Frankfort, Kentucky.

Rosemond, A.D., S.R. Reice, J.W. Elwood, and P.J. Mulholland. 1992. The effects of stream acidity on benthic invertebrate communities in the south-eastern United States. Freshwater Biology 27: 193-209.

Seaber, P.R., F.P. Kapinos, and G.L. Knapp. 1987. Hydrologic unit maps. United States Geological Survey Water-Supply Paper 2294.
Shannon, C.E. 1948. A mathematical theory of communication. Bell System Technical Journal 27: 379-423 and 623-656.

Simpson, K.W., R.W. Bode, and J.R. Colquhoun. 1985. The macroinvertebrate fauna of an acid-stressed headwater stream system in the Adirondack Mountains, New York. Freshwater Biology 15: 671-681.

Stoddard, J.L., D.P. Larsen, C.P. Hawkins, R.K. Johnson, and R.H. Norris. 2006. Setting expectations for the ecological condition of streams: the concept of reference condition. Ecological Applications 16(4): 1267-1276.

Stribling, J.B., B.K. Jessup, J.S. White, D. Boward, and M. Hurd. 1998. Development of a benthic index of biotic integrity for Maryland streams. Chesapeake Bay and Watershed Programs, Monitoring, and Non-Tidal Assessment Division. CBW-MANTA – EA-98-3.

Sutcliffe, D.W. and A.G. Hildrew. 1989. Invertebrate communities in acid streams. In: Acid toxicity and aquatic animals. Society for Experimental Biology Seminar Series, Volume 34. Morris, R., E.W. Taylor, D.J.A. Brown, and J.A. Brown, eds. Cambridge University Press. pp. 13-30.

Unites States Code. Accessed June 24, 2011. http://www4.law.cornell.edu/uscode/.

United States Environmental Protection Agency. 1990. Freshwater macroinvertebrate species list including tolerance values and functional feeding group designations for use in rapid bioassessment protocols. Report Number 11075.05.

Van Sickle, J. 1998. A set of programs for mean similarity analysis. United States Environmental Protection Agency National Health and Environmental Effects Laboratory, Western Ecology Division. Corvallis, Oregon, USA.

Van Sickle, J. and R.M. Hughes. 2000. Classification strengths of ecoregions, catchments, and geographic clusters for aquatic vertebrates in Oregon. Journal of the North American Benthological Society 19: 370-384.

Vannote, R. L., G. W. Minshall, K. W. Cummins, J. R. Sedell, AND C. E. Cushing. 1980. The river continuum concept. Canadian Journal of Fisheries and Aquatic Sciences 37: 130-137.

Whittaker, R.H. 1962. Classification of natural communities. Botanical Review 28: 1-239.

Yoder, C.O. and E.T. Rankin. 1995. Biological criteria program development and implementation in Ohio. Chapter 9 in *Biological assessment and criteria: tools for water resource planning and decision making*, W.S. Davis and T.P. Simon, eds. (pp. 109 – 144). CRC Press, Boca Raton.

1. Habitat Assessments

The Department has adopted the habitat assessment methods outlined in USEPA's Rapid Bioassessment Protocols (RBP; Plafkin, et al. 1989) and subsequently modified[1]. The matrix used to assess habitat quality is based on key physical characteristics of the water body and surrounding lands. All parameters evaluated represent potential limitations to the quality and quantity of instream habitat available to aquatic biota. These, in turn, affect community structure and composition.

The main purpose of the habitat assessment is to account for the limitations that are due to existing stream conditions. This is particularly important in cause/effect and cumulative impact studies where the benthic community at any given station may already be self-limited by background watershed and habitat conditions or impacts from current land uses. In order to minimize the effects of habitat variability, every effort is made to sample similar habitats at all stations. The habitat assessment process involves rating twelve[1] parameters as excellent, good, fair, or poor, by assigning a numeric value (ranging from 20 - 0[1]), based on the criteria included on the Habitat Assessment Field Data Sheets.

The twelve habitat assessment parameters used in the DEP-RBP evaluations for Riffle/Run prevalent (and Glide/Pool prevalent) streams are discussed below. The Glide/Pool parameters that differ from the Riffle/Run parameters are shown in italics. The first four parameters evaluate stream conditions in the immediate vicinity of the benthic macroinvertebrate sampling point:

- **Instream Fish Cover** - evaluates the percent makeup of the substrate (boulders, cobble, other rock material) and submerged objects (logs, undercut banks) that provide refuge for fish.

- **Epifaunal Substrate** - evaluates riffle quality, i.e. areal extent relative to stream width and dominant substrate materials that are present. *(In the absence of well-defined riffles, this parameter evaluates whatever substrate is available for aquatic invertebrate colonization.)*

- **Embeddedness** - estimates the percent (vertical depth) of the substrate interstitial spaces filled with fine sediments. *(**pool substrate characterization:** evaluates the dominant type of substrate materials, i.e. gravel, mud, root mats, etc. that are more commonly found in glide/pool habitats.)*

1. Plafkin et al. (1989) originally presented nine habitat assessment parameters divided into three different scoring ranges of 20-0, 15-0, and 10-0. Modifications to these original habitat methods were presented at several seminars following this 1989 publication. These modifications added one more habitat parameter to each of the three original categories; bringing the total parameters to 12. The scoring ranges eventually were increased to 20-0 for all 12. This Habitat Protocol has undergone several more iterations – resulting in yet more variations from the original and the Department's current 12 criteria - 20 point scoring habitat assessment method.

- **Velocity/Depth Regime** - evaluates the presence/absence of four velocity/depth regimes - fast-deep, fast-shallow, slow-deep, and slow-shallow. (Generally, shallow is <0.5m and slow is <0.3m/sec. ***Pool variability:*** *describes the presence and dominance of several pool depth regimes.*)

The next four parameters evaluate a larger area surrounding the sampled riffle. As a rule of thumb, this expanded area is the stream length defined by how far upstream and downstream the investigator can see from the sample point.

- **Channel Alteration** - primarily evaluates the extent of channelization or dredging but can include any other forms of channel disruptions that would be detrimental to the habitat.

- **Sediment Deposition** - estimates the extent of sediment effects in the formation of islands, point bars, and pool deposition.

- **Riffle Frequency (pool/riffle or run/bend ratio)** - estimates the frequency of riffle occurrence based on stream width. (***Channel sinuosity:*** *the degree of sinuosity to total length of the study segment.)*

- **Channel Flow Status** - estimates the areal extent of exposed substrates due to water level or flow conditions.

The next four parameters evaluate an even greater area. This area is usually defined as the length of stream that was electro-shocked for fish (or an approximate 100 meter stream reach when no fish were sampled). It can also take into consideration upstream land-use activities in the watershed:

- **Condition of Banks** - evaluates the extent of bank failure or signs of erosion.

- **Bank Vegetative Protection** - estimates the extent of stream bank that is covered by plant growth providing stability through well-developed root systems.

- **Grazing or Other Disruptive Pressures** - evaluates disruptions to surrounding land vegetation due to common human activities, such as crop harvesting, lawn care, excavations, fill, construction projects, and other intrusive activities.

- **Riparian Vegetative Zone Width** - estimates the width of protective buffer strips or riparian zones. This is a rating of the buffer strip with the least width.

It is best to conduct the habitat assessment after sampling since the investigator has observed all conditions in the sampled segment and immediate surrounding watershed. After all parameters in the matrix are evaluated and scored, the scores are summed to derive a habitat score for that station. The “optimal” category scores range from 240-192; “sub-optimal” from 180-132; “marginal” from 120-72; and “poor” is 60 or less. The gaps between these categories are left to the discretion of the investigator’s best professional judgment.

2. Benthic Macroinvertebrates

2.A. Net Mesh Considerations

In recent years, many state water quality programs, federal agencies (e.g. USEPA, USGS), and other water quality monitoring organizations began using net sampling devices with 500µ □mesh nets. In order to conform to this trend, the 500µ □net mesh size has been adopted for the Department's D-frame sampler used in the DEP-RBP sampling method (described below). Future references to the D-frame sampler in the document assume 500-µ mesh netting. The net mesh size of other screen samplers has not changed and still is to be 800-900 µ. Because of this net mesh size change, the mesh size of the sampler used must be noted on field and bench identification sheets for the collected benthic sample.

2.B. Qualitative Methods

The type of sampling gear used is dependent on survey type and site-specific conditions. The recommended gear in wadeable streams are 3' x 3' flexible kick-screens and 12-inch diameter round D-frame nets. In larger streams or rivers, grab-type samplers may be used to obtain qualitative samples. While generally thought of as quantitative devices, Eckman, Peterson, or Petite Ponar grab samplers can also be used to obtain qualitative data. The type of gear, dimensions, and mesh size must be reported for all collections. When more than one gear type is used, the results must be recorded separately.

Physical variables should be matched as closely as possible between background and impact stations when selecting locations for placement of the sampling gear within each station. Matching these variables helps minimize or eliminate the effects of compounding variables.

Macrobenthos often exhibit clustered distributions, and if the sampling points are selected in close proximity to each other, a single clustered population may be obtained rather than a generalized measure of the overall population within the selected sub-habitat. Spacing the sampling points as far apart as possible within the sub-habitat can minimize the problem of clustered distributions.

2.B.1. Kick-screen. A common qualitative sampling method uses a simple hand-held kick-screen. This device is designed to be used by two persons. However, with experience, it may be used by one person and still provide adequate results. The kick-screen is constructed with a 3' x 3' piece of net material (800-900 µ mesh size) fastened to two dowel handles (approximately 1"d. X 4' long).

2.B.1.a. Traditional Method. Facing up stream, one person places the net in the stream with the bottom edge of the net held firmly against the streambed. An assistant then vigorously kicks the substrate within a 3' x 3' area immediately upstream of the net to a depth of 3" - 4" (approximately 10 cm). The functional depth sampled may vary due to ease of disturbance as influenced by substrate embeddedness.

The amount of effort expended in collecting each sample should be approximately equivalent in order to make valid comparisons. The effort, expressed as area, must be reported for all collections.

Collect a minimum of four screens at each site. Initial sampling should be conducted in riffle areas. Collection in additional habitats to generate a more complete taxa list can be conducted at the discretion of the investigator. Initial analysis of the data must be limited to the riffle data for standardization. A second analysis including other habitats may be conducted as needed.

Data observations shall be recorded on a standard field sheet created for each station sampled. Record the relative abundance of each recognizable Family in each individual collection in the field. Relative abundance categories, with the observed "total" ranges indicated in parenthesis include: rare (0-3), present (3-10), common (11-24), abundant (25-99), and (occasionally) very abundant (100+). The investigator, at his/her discretion, may elect to enumerate certain target taxa.

Recording the results of each collection has several advantages that are lost if the data are composited for each station:

a. A stressed or enriched community often exhibits little variability in community structure over an area while a healthy community should have a more complex structure. If varied taxa are found on each screen, the community is probably complex, while the presence of only a few dominant taxa on every screen indicates the community is a simple one.

b. Collecting intolerant taxa in a majority of screens is a good indication of an unstressed community. However, collecting intolerant taxa in only one out of four screens may be an indication that the intolerant taxa have only a marginal existence at that location. A comparison of the composited taxa lists for each location may not indicate the rarity of the intolerant taxa, but this rarity would be readily apparent if the taxa lists for individual screens were compared.

c. Separate screen taxa lists provide information concerning the distribution of taxa. For example, mayflies are taken in one of four screens at the background station and in none of the four screens at the impact station. All the other taxa collected at both the stations are tolerant forms. Based on a composited taxa list for each station, one might conclude that the impact station is depressed due to the absence of mayflies. However, the individual screen taxa lists would indicate that the mayflies may have a clumped distribution and there is a possibility that the collector simply missed the clumps at the impact station. This will be apparent to the biologist while in the field and he/she can continue collecting until comfortable that mayflies are indeed absent or less abundant at the impact station. Later, it can be reported, for example, that 4 of 10 screens contained mayflies at the background station while only 1 of 10 screens contained mayflies at the impact station. This is an instance when the collector, while still in the field, may choose to count the mayflies in each screen (especially if the background screens had many mayflies while the impact screens only had one or two).

d. Separate screen data can lend weight to an analysis when classification techniques (ordination or clustering) are used. Results that cluster or score the individual background screens differently than the individual impact screens indicates a difference between the locations. When the classification technique scores background and impact screens in an apparent random manner, then it is likely that there is no impact or that the natural variability is large and masks any impacts.

Individuals of representative taxa for a station may be composited in a single vial and preserved for later laboratory verification or identification. Generally, the level of taxonomic identification would follow that as listed in section 2.E.1.

Answers to several questions can be useful in subsequent analysis and can be stored with the taxa lists as remark fields. The answers to the following questions, which require collector judgment, can be recorded in the field on a coded form. What are the dominant and rare taxa? Are there any taxa that are found to be unusually abundant?

2.B.1.b Assessment Method. This method is used for assessments conducted as part of the Statewide Surface Waters Assessment Program and employs the same kick screen gear, physical disturbance techniques, and relative abundance determinations as the traditional method (2.B.1.a). The main difference is that only two kicks are usually required and macroinvertebrate identifications are done streamside to family level taxonomy with hand-held lens (10X) if necessary. Data are recorded on standard field forms. Refer to the Statewide Surface Waters Assessment Protocol for further details.

2.B.2. D-Frame. The handheld D-frame sampler consists of a bag net attached to a half-circle ("D" shaped) frame that is 1' wide. The net's design is that of an extended, round bottomed bag (500µ mesh size). The methodology is basically the same as with the kick-screen - except for the following points: one person, facing downstream and holding the net firmly on the stream bottom, employs the net. One "**D-frame effort**" is defined as such: the investigator vigorously kicks an approximate area of 1 m^2 immediately upstream of the net to a depth of 10 cm (or approximately 4", as the embeddedness of the substrate will allow) for approximately one minute. All benthic dislodgement and substrate scrubbing should be done by kicks only. Substrate handling should be limited to only moving large rocks or debris (as needed) with no hand washing. Since the width of the kick area is wider than the net opening, net placement is critical in order to assure all kicked material flows toward the net. Avoiding areas with crosscurrents, the substrate material from within the square meter area should be kicked toward the center of the area – above the net opening.

The concepts and field forms concerning field recording of invertebrate data discussed in the kick-screen method section (2.B.1a) also apply to the D-frame method.

2.C. Semi-Quantitative Method (DEP-RBP):

In Plafkin (1989), USEPA presented field-sampling methods designed to assess impacts normally associated with pollution impacts, cause/effect issues, and other water quality degradation problems in a relatively rapid manner. These are referred to as Rapid Bioassessment Protocols (RBPs). The DEP-RBP method is a bioassessment technique involving systematic field collection and subsequent lab analysis to allow detection of benthic community differences between reference (or control) waters and waters under evaluation. The DEP-RBP is a modification of the USEPA RBP III (Plafkin, et al; 1989); designed to be compatible with Pennsylvania's historical database. Modifications include: 1) the use of a D-frame net for the collection of the riffle/run samples, 2) different laboratory sorting procedures, 3) elimination of the CPOM (coarse particulate organic matter) sampling, and 4) metrics substitutions. Unlike the USEPA's RBP III methodology, no field sorting is done. Only larger rocks, detritus, and other debris are rinsed and removed while in the field before the sample is preserved. While USEPA's RBP III method was designed to

compare impacted waters to reference conditions (cause/effect approach), the DEP-RBP modifications were designed for un-impacted waters, as well as impacted waters.

2.C.1. Sample Collection. The purpose of the standardized DEP-RBP collection procedure is to obtain representative macroinvertebrate fauna samples from comparable stations. The DEP-RBP assumes the riffle/run habitat to be the most productive habitat. Riffle/run habitats are sampled using the D-frame net method described above. The number of D-frame efforts is dependent on the type of survey conducted as described below:

2.C.1.a. Limestone Streams. For limestone stream surveys, two paired D-frame efforts are collected from each station - one from an area of fast current velocity and one from an area of slower current velocity within the same riffle.

2.C.1.b. Antidegradation Surveys. For Antidegradation surveys, it is necessary to characterize macroinvertebrate fauna communities from an area larger than a single riffle. Therefore, an Antidegradation survey station is defined as a stream reach of approximately 100 meters in length. At each station, six "D-frame efforts" are collected. Make an effort to spread the samples out over the entire reach. Choose the best riffle habitat areas and be certain to include areas of different depths (fast and slow) and substrate types that are typical of the riffle.

The resulting "D-frame efforts" (six for Anti-degradation, two for other survey types) are composited into one sample jar (or more as necessary). Care must be taken to minimize "wear and tear" on the collected organisms when compositing the materials. It is recommended that the benthic material be placed in a bucket and filled with water to facilitate gentle stirring and mixing. The sample is preserved in ethanol and returned to the lab for processing.

2.C.2. Sample Processing. Samples collected with a D-frame net are generally considered to be qualitative. However, the preserved samples can be processed in a manner which yields data that are "semi-quantitative" - data that were collected by qualitative methods but gives information that is almost statistically as strong as that collected by quantitative methods.

The following procedure is adapted from USEPA 1999 RBP methodology and used to process qualitative D-frame samples so that the resulting data can be analyzed using benthic macroinvertebrate biometric indices (or "metrics"). Equipment needed for the benthic sample processing are:

- 2 large laboratory pans gridded into 28 squares***** (more gridded pans may be necessary depending on the size of the sample);
- an illuminated magnifying viewer;
- slips of paper (numbered from 1 to 28) for drawing random numbers;
- forceps (or any tools that can be used to pick floating benthic organisms); and
- grid cutters made from tubular material that approximates an inside area of 4 in^2*****.

***** *USEPA's (1989) gridding techniques suggest using "5 cm x 5 cm" (2" x 2") grids. Existing equipment consisted of 14" x 8" x 2" pans which were conducive to dividing into 2" x 2" grids and thus, contained 28 squares. The 4-in^2 grid cutters conform to these pan dimensions. While pan size is not critical, the number of grids (28) must be maintained if any basic density comparisons wish to be made between samples. Grid cutters (or similar sub-sampling devices) used with different sized pans should conform to the pans' grid dimensions.*

The procedure described below begins with the premise that the collected samples have been properly composited according to the type of survey. For Antidegradation surveys, a station sample represents a composition of six D-frame efforts (collected from fast and slow riffle areas in a 100 meter reach). For Limestone surveys, a station sample is a composition of two D-frame efforts.

Following the steps listed below; process each composited D-frame sample to render a sub-sample size targeted for the specific survey type. The targeted sub-sample size for Antidegradation surveys is 200 benthic organisms and 300 for Limestone surveys (± 20% for each).

a. The composited sample is placed in a 28-square gridded pan (Pan1). It is recommended that the sample be rinsed in a standard USGS No. 35 sieve (or sieve bucket) to remove fine materials and residual preservative prior to sub-sampling.

b. The sample is gently stirred to disperse the contents evenly throughout Pan1 as thoroughly as possible. (In order to ease mixing and to minimize "wear-and-tear" on the more delicate organisms, water may be added to the pan to the depth of the sample material before stirring.)

c. Randomly select a grid using the 28 random number set and, using the grid cutters, remove the debris and organisms entirely from within the grid cutter (centered over the selected grid and "cut" into the debris) and place removed materials in a second gridded pan (Pan2).

    i. Float and pick, count, and sub-total all identifiable organisms (excluding pupae, larval bodies missing too many critical structures to render confident IDs, extremely small instar larvae, empty shells or cases, and non-benthic taxa) from each cut grid placed in Pan2. Repeat until at least 4 grids have been sub-sampled from Pan1. If, after 4 Pan1 grids have been sorted, the sub-total is less than the targeted sub-sample (20 ± 20%), then continue to remove and sort grids one at a time until 200 organisms (± 20%) are obtained from Pan2. If the benthic organism yield from the 4 Pan1 grids exceeds the 200 ± 20% target (240+), then proceed to Step ii.

    ii. With all of the 240+ identifiable organisms remaining in Pan2, randomly select one grid and "back count" (removing) all the organisms from that grid. Repeat one grid at a time until the bug count remaining in Pan2 satisfies the "200 ± 20%" rule.

d. If not identified immediately, the sub-sample should be preserved and properly labeled for future identification.

e. The benthic material remaining (Pan1) after the target sub-sample has been picked can be returned to its original sample jar and preserved. They shall be retained in accordance with QA retention times as specified for the respective survey type.

f. Any grid chosen must be picked in its entirety.

g. Record the final grid counts selected for each gridding phase (Pan1, Pan2, and Pan2 "back counting" as necessary) on the lab bench ID sheet for the sample.

Processing larger, excessive amounts of D-frame sample debris

Hopefully, the collector will rarely have very large amounts of D-frame materials to process. The reduction of large materials by careful removal, inspection, and rinsing in a bucket or using a sieve prior to field preservation or at the lab is encouraged. However, if the amount of material composited in the field jars exceeds the functional sorting capacity of Pan1, then follow this guidance:

- Evenly distribute the material between as many pans as necessary.
- From each pan (Pan1a, Pan1b, etc.), remove debris and organisms from 4 random grids and place in Pan2 as described in Step 2.C.2.c above.
- Once the required 4 grids from each Pan1 have been placed in Pan2, evenly and gently redistribute the materials as in Step 2.C.2.b.
- Then, resume processing, again as described in Step 2.C.2.c, selecting a grid from Pan2 and placing the materials into a gridded Pan3.
- Process this material and repeat as described in Step 2.C.2.c.i until the targeted 200 ± 20% sub-sample is obtained from Pan3.
- If, after processing 4 grids, the +20% upper limit (240+) is obtained, follow "back counting" method in Step 2.C.2.c.ii.
- Once the targeted sub-sample is reached, continue with Step 2.C.2.d.

2.D. Identification

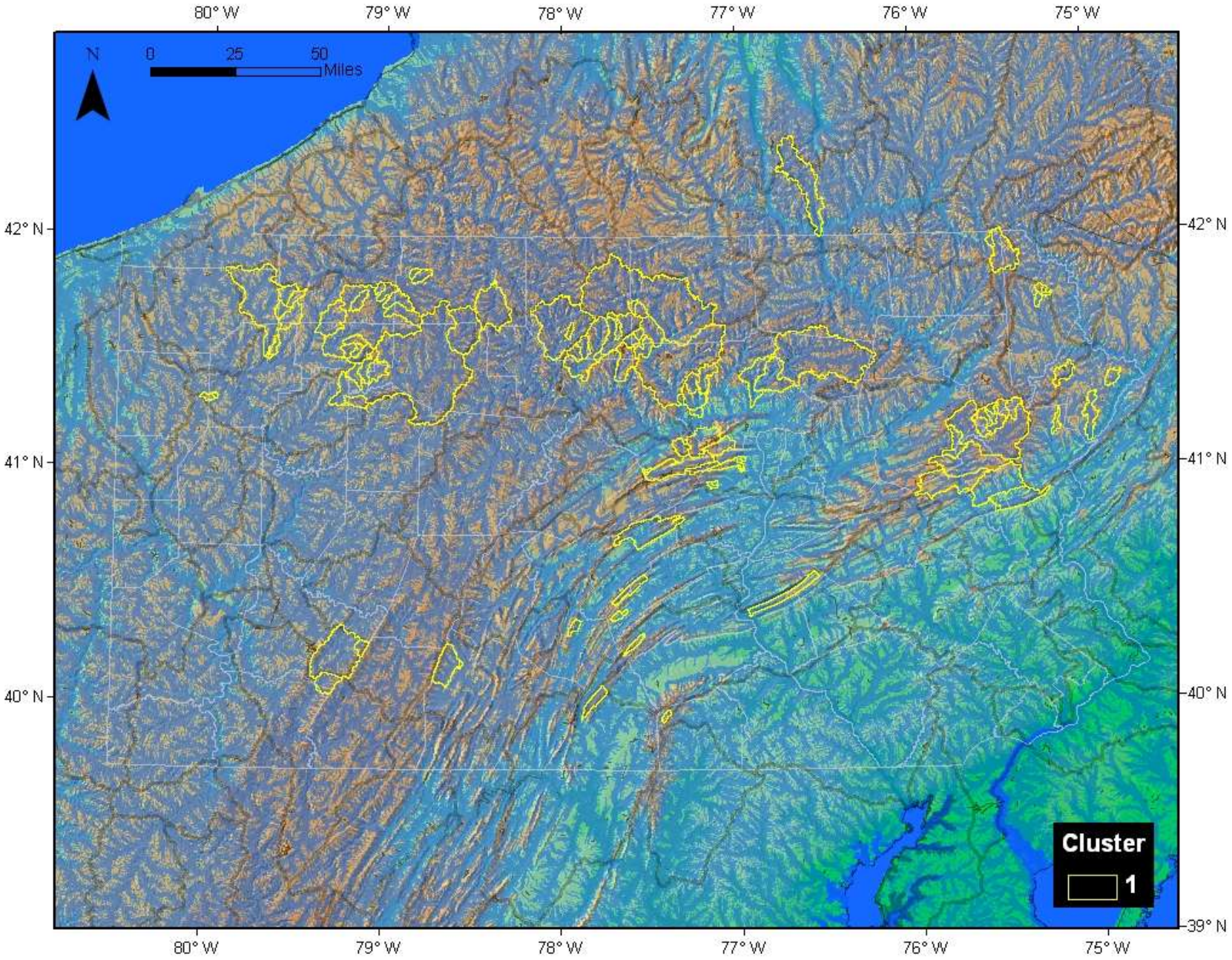
2.D.1. Taxonomic Level. The level of identification for most aquatic macroinvertebrates will be to genus. Presently, the identification of Chironomidae, or midges, is to the family level. Some individuals collected will be immature and not exhibit the characteristics necessary for confident identification. Therefore, the lowest level of taxonomy attainable will be sufficient. Certain groups, however, may be identified to a higher taxonomic level as follows:

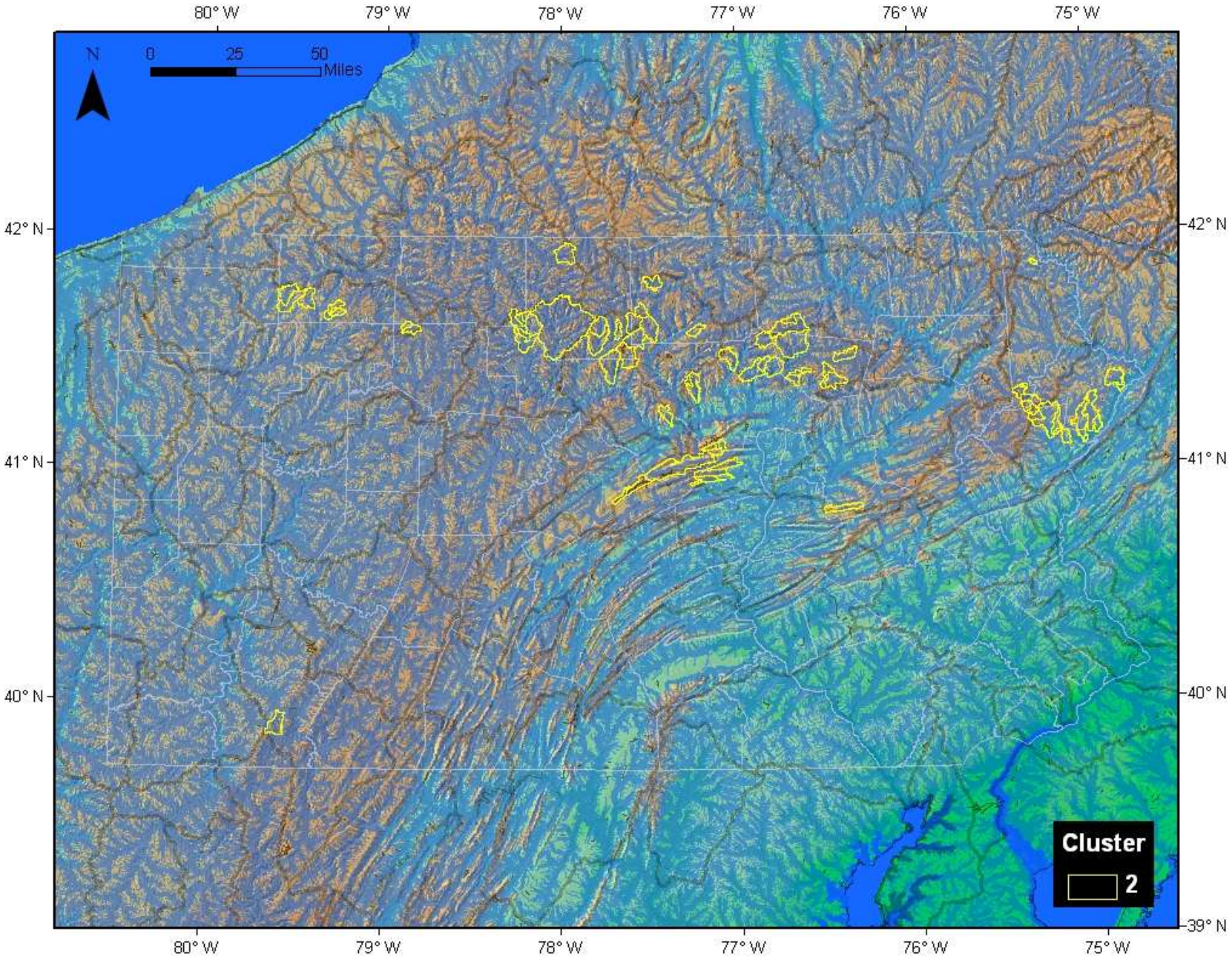
Snails (Gastropoda) - Family
Clams, mussels (Bivalvia) - Family
Flatworms (Turbellaria)
    identifiable planariids - genus
    or Family Planariidae
    others – Class Turbellaria
Segmented worms (Annelida)
    aquatic earthworms & tubificids - Class Oligochaeta
    leeches - Class Hirudinea
    Moss animacules - Phylum Bryozoa
Proboscis worms – Phylum Nemertea
Roundworms - Phylum Nematoda
Water mites- "Hydracarina" (an artificial taxonomic grouping of several mite superfamilies)

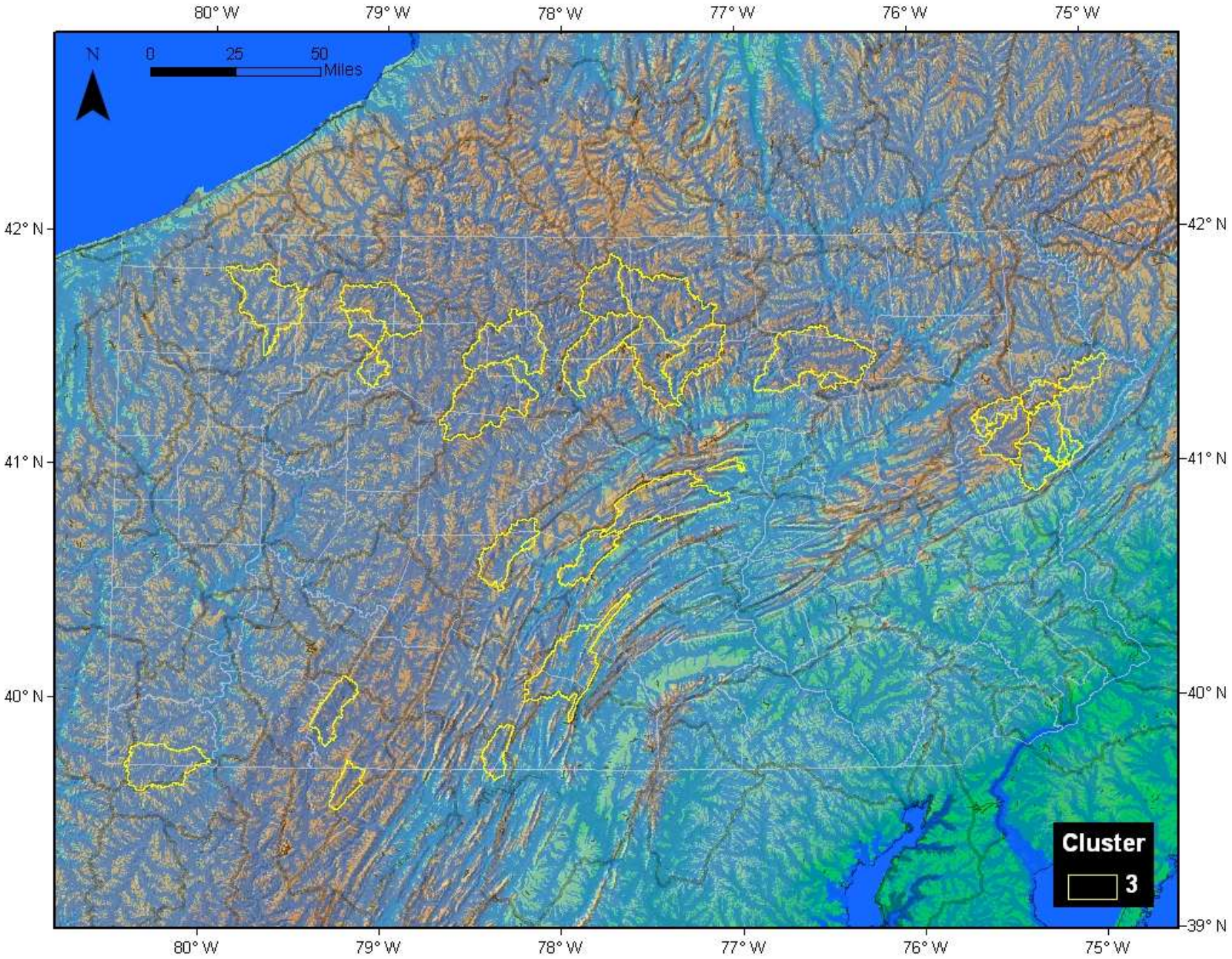
2.D.2. Verifications. For Quality Assurance purposes, certain laboratory invertebrate processing procedures should be checked routinely. Normally, a colleague may perform these spot checks. These include the floating/picking steps, taxonomic identifications, and total taxa list scans:

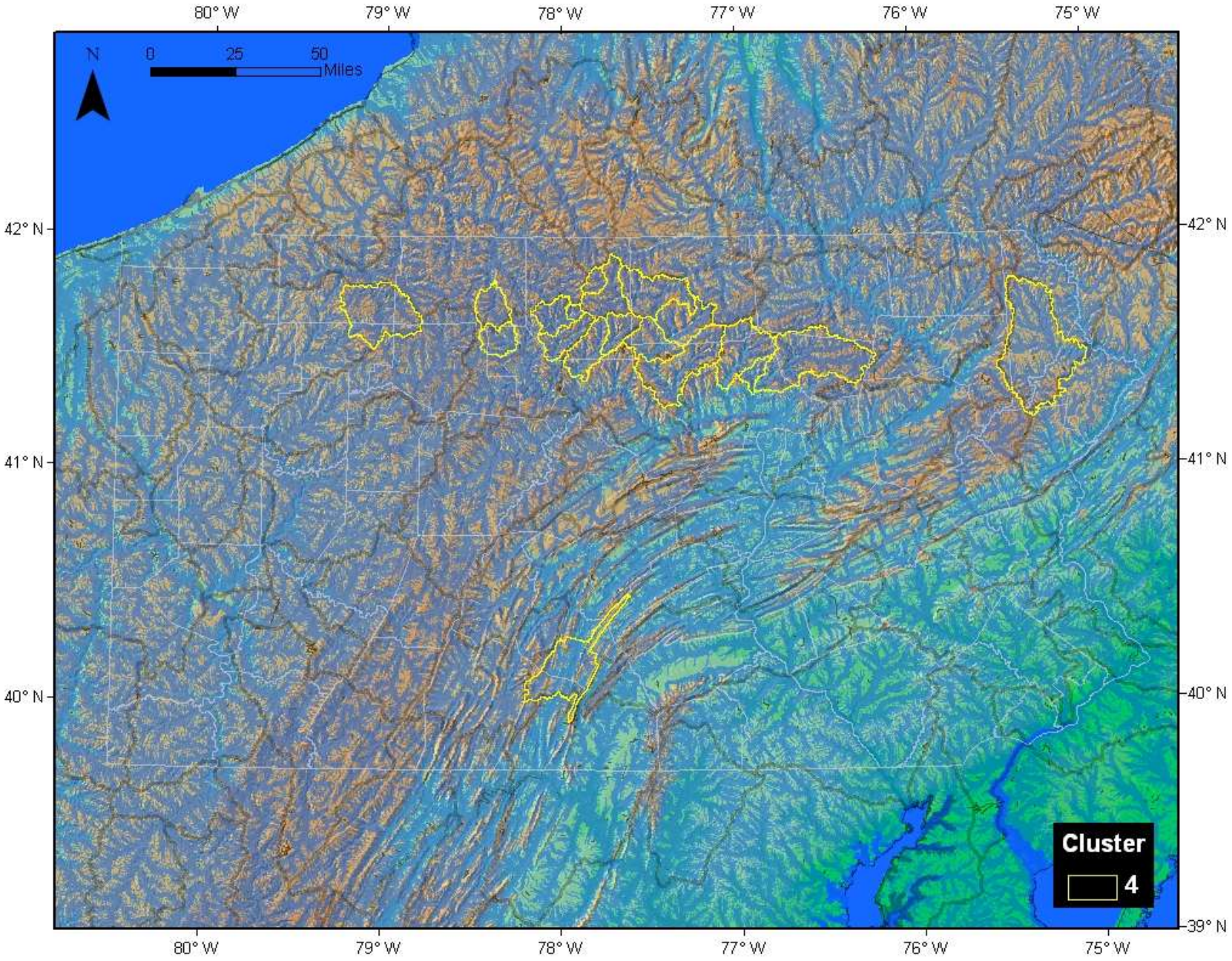
a. Sorting. After the floating and picking has been completed for samples that require this treatment (Pa-RBP, Surber-type, multi-plate, and grab samples), the residue should be briefly scanned before discarding to assure that the sample has been sufficiently "picked". This should be done for 10% of the samples (or at least one sample) per survey.

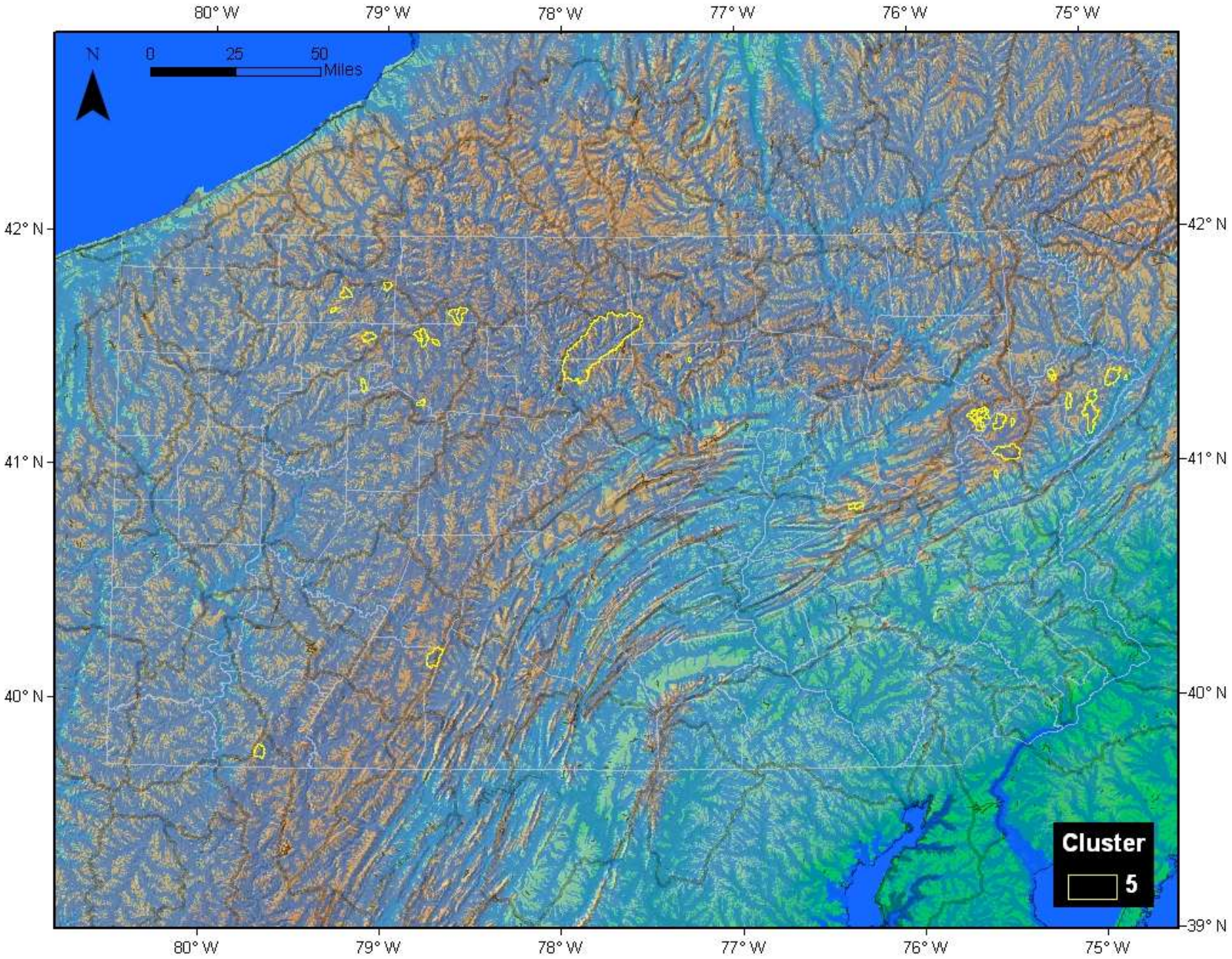
b. Identification. For samples not involving litigation or enforcement issues, laboratory bench ID sheets for all samples should be reviewed. Any unusual taxa or taxa that are not typical to the type of stream or water quality condition that was surveyed, should be checked. For samples involving legal issues, representative specimens of each taxon may need to be verified by independent expert taxonomists.

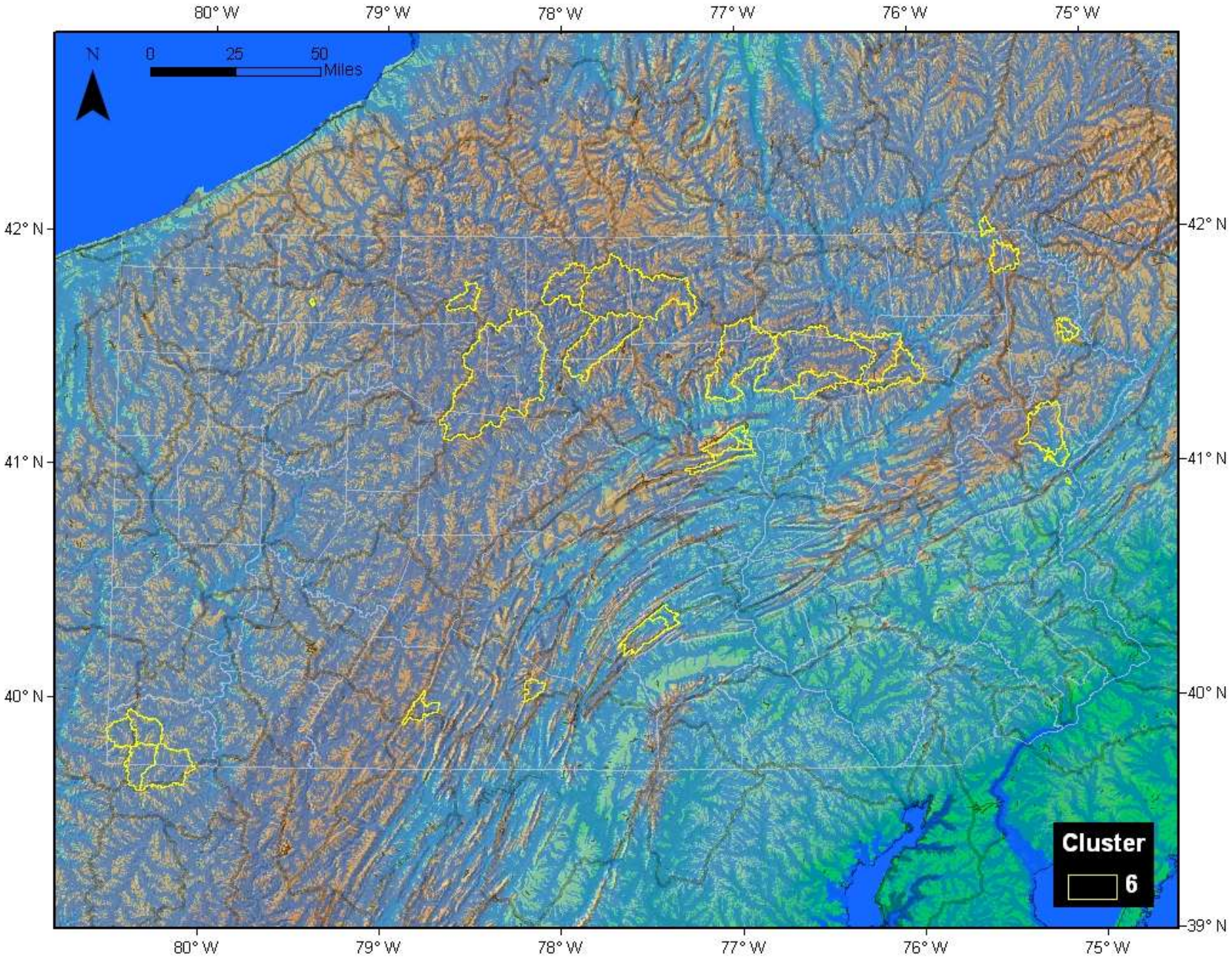

80° W
79° W
78° W
77° W
76° W
75° W
N
0
25
50
Miles
42° N
41° N
40° N
39° N
Cluster
1

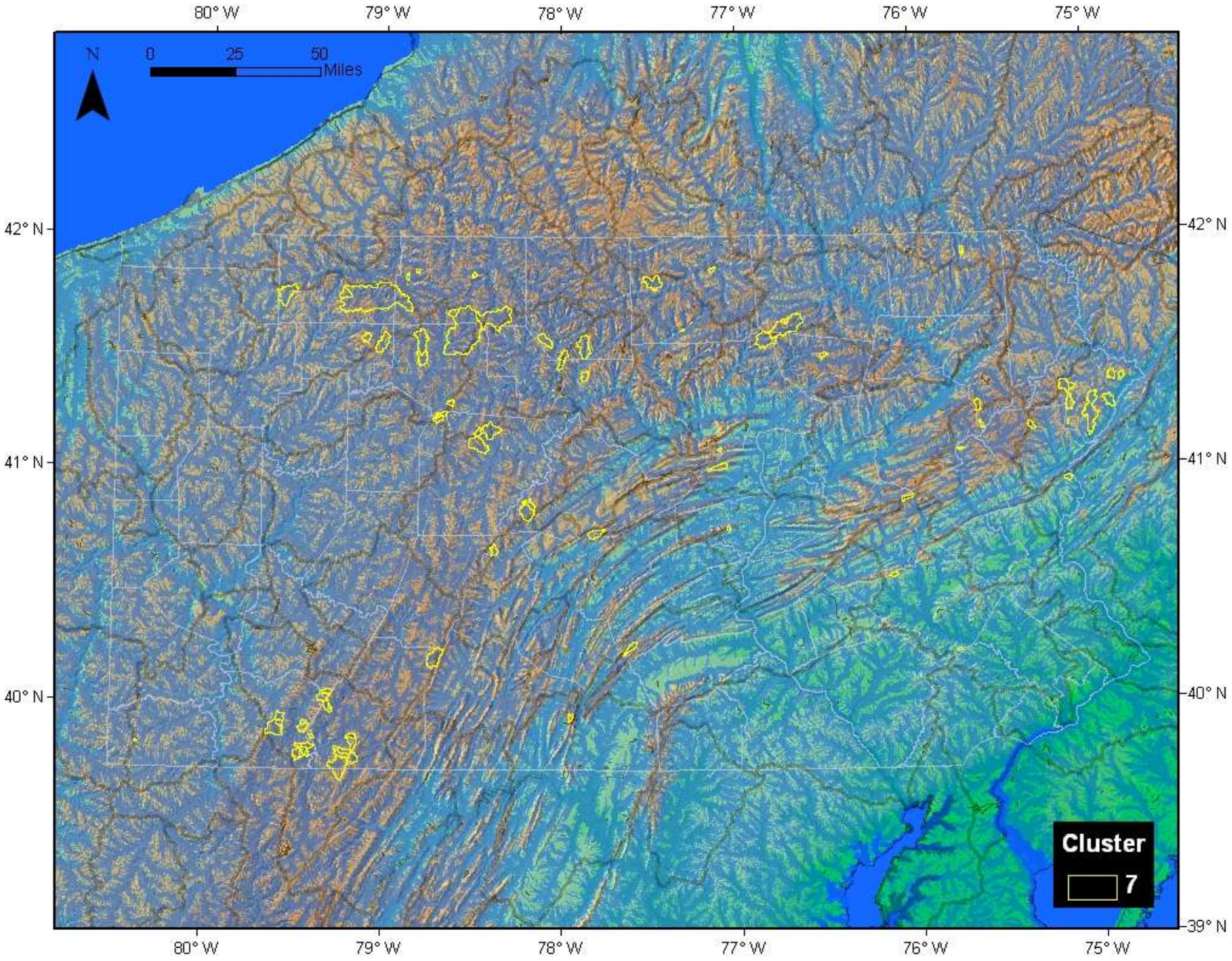

80° W
79° W
78° W
77° W
76° W
75° W
N
0
25
50
Miles
42° N
41° N
40° N
39° N
Cluster
2

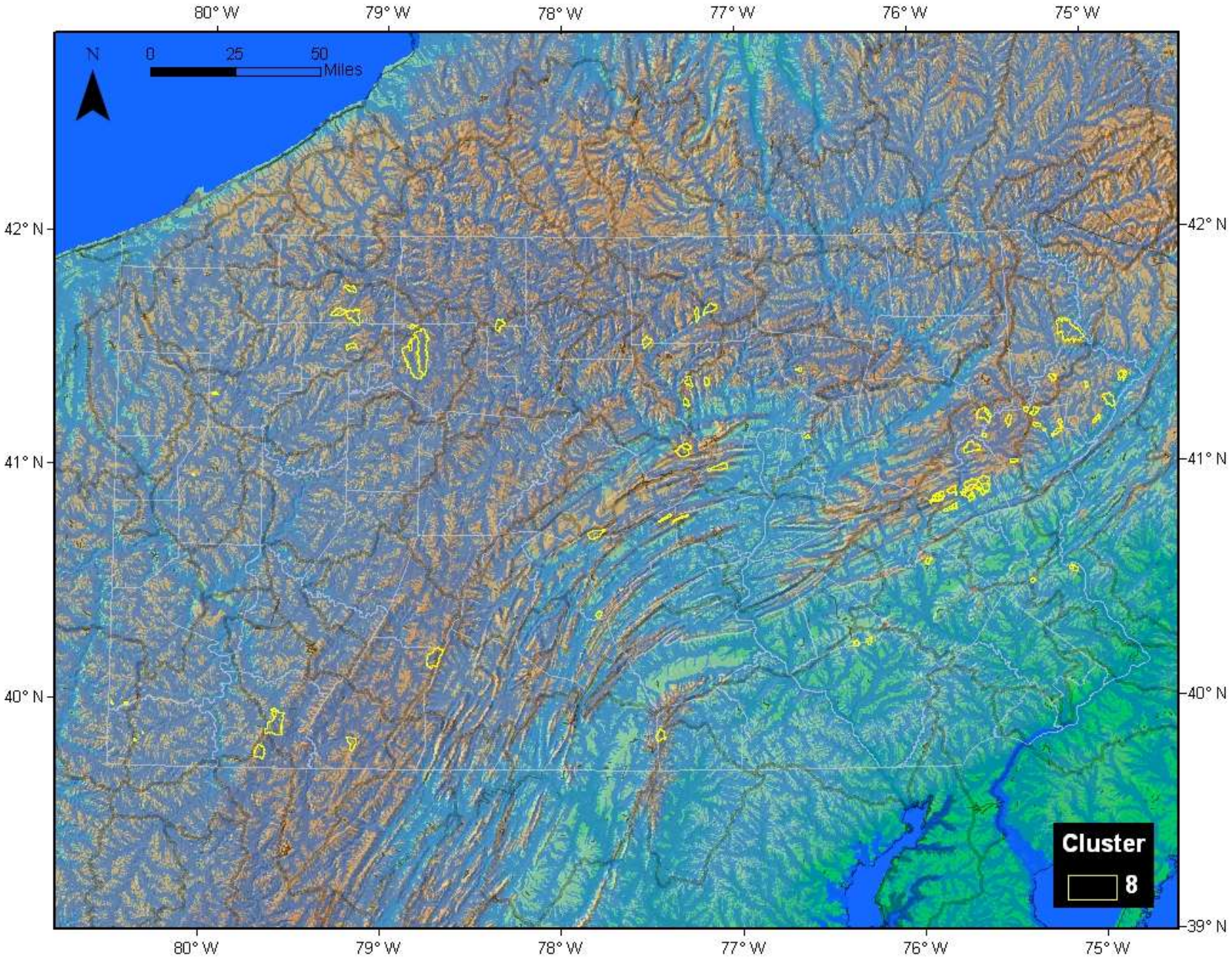

80° W
79° W
78° W
77° W
76° W
75° W
N
0
25
50
Miles
42° N
41° N
40° N
39° N
Cluster
3

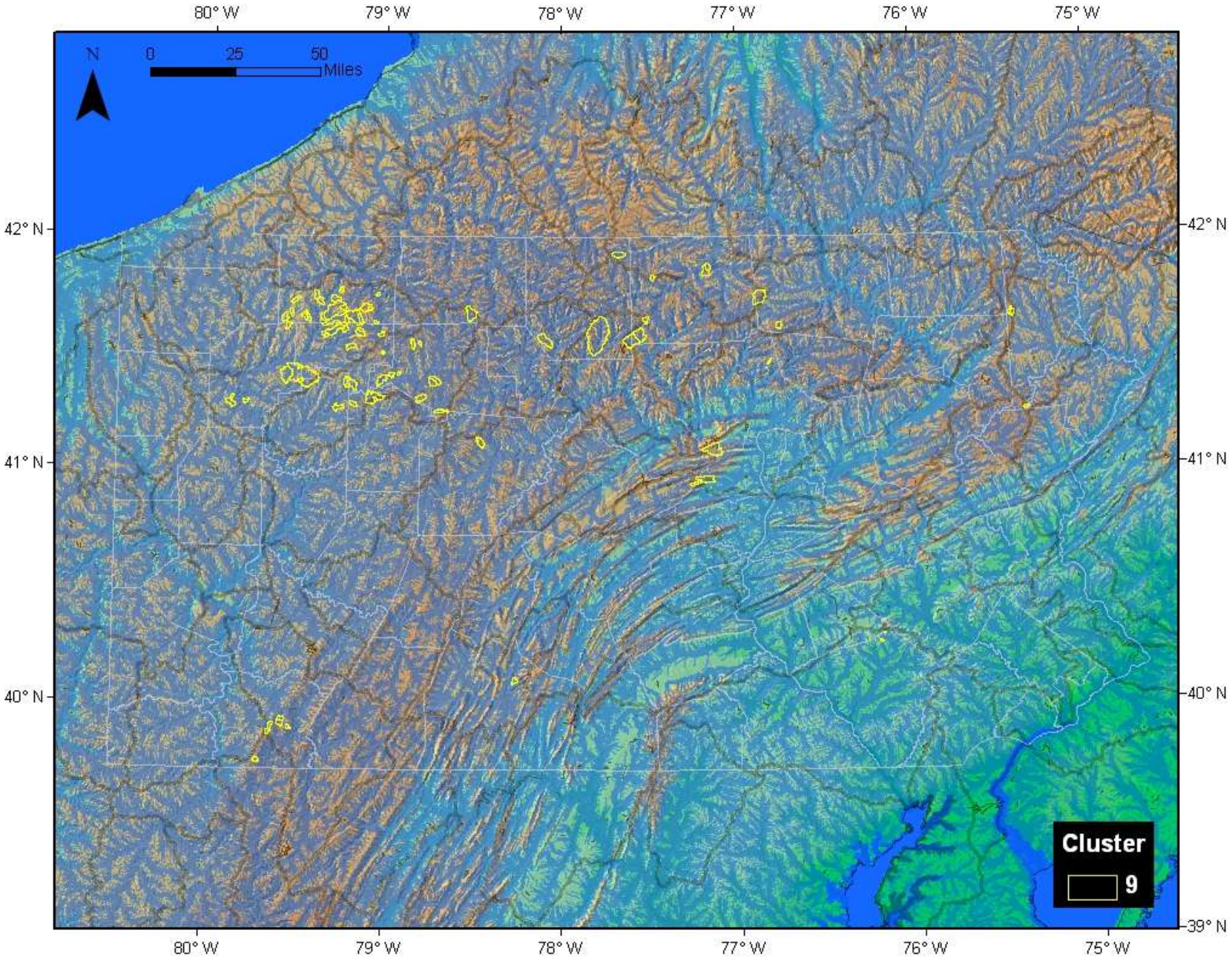

80° W
79° W
78° W
77° W
76° W
75° W
N
0
25
50
Miles
42° N
41° N
40° N
39° N
Cluster
4

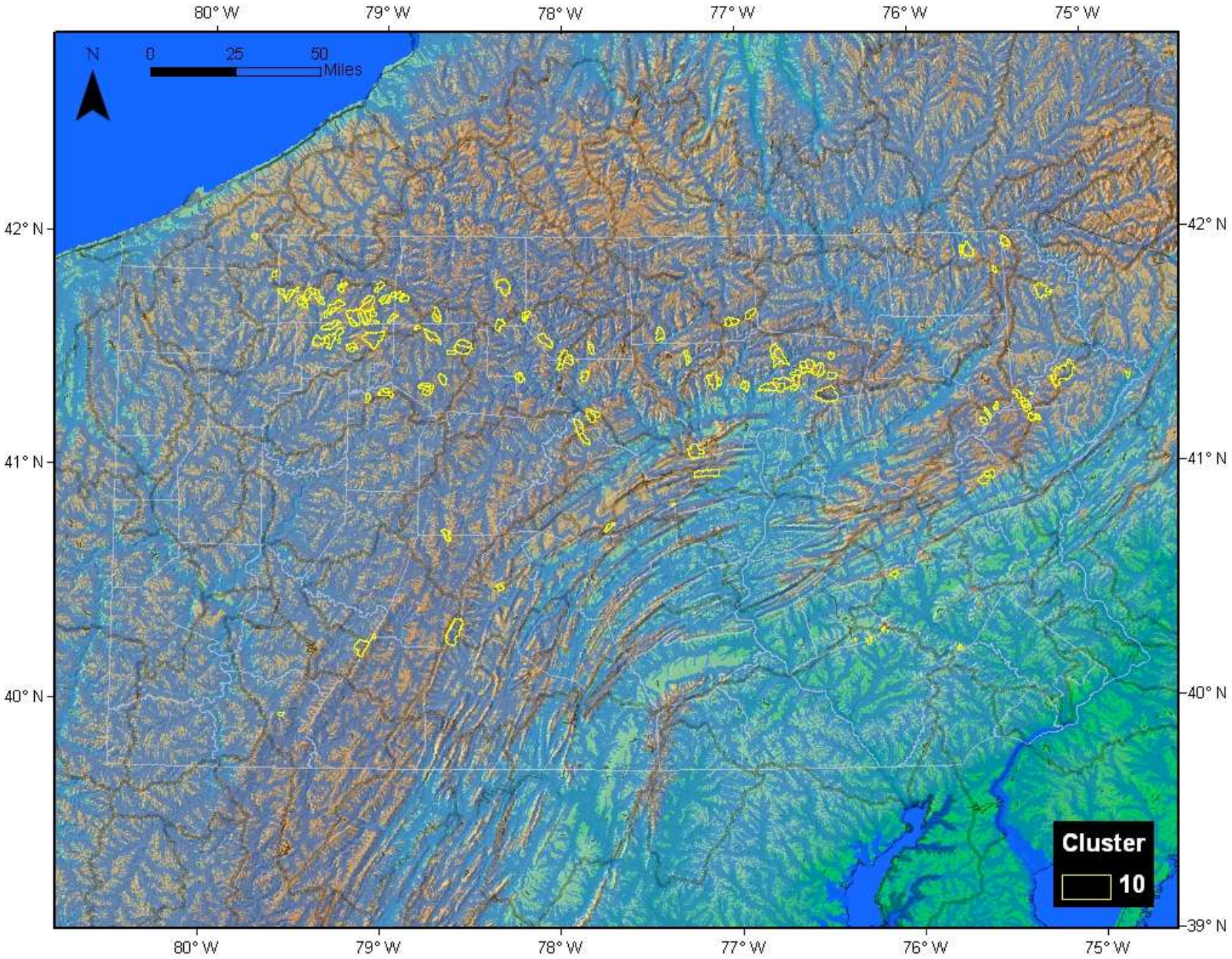

80° W
79° W
78° W
77° W
76° W
75° W
N
0
25
50
Miles
42° N
41° N
40° N
39° N
Cluster
5

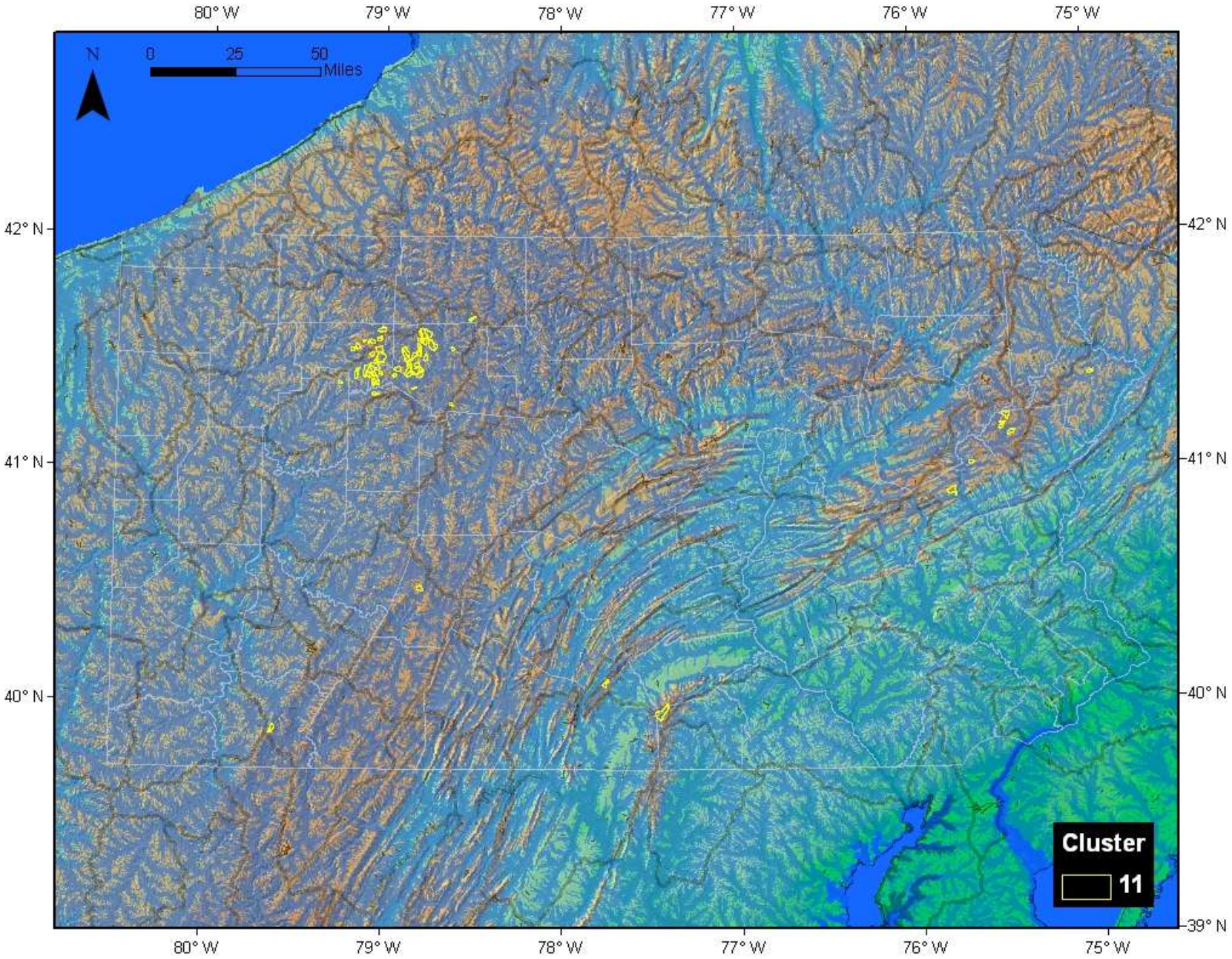

80° W
79° W
78° W
77° W
76° W
75° W
N
0
25
50
Miles
42° N
41° N
40° N
39° N
Cluster
6


80° W
79° W
78° W
77° W
76° W
75° W
N
0
25
50
Miles
42° N
41° N
40° N
39° N
Cluster
7


80° W
79° W
78° W
77° W
76° W
75° W
N
0
25
50
Miles
42° N
41° N
40° N
39° N
Cluster
8


80° W
79° W
78° W
77° W
76° W
75° W
N
0
25
50
Miles
42° N
41° N
40° N
39° N
Cluster
9


80° W
79° W
78° W
77° W
76° W
75° W
N
0
25
50
Miles
42° N
41° N
40° N
39° N
Cluster
10


80° W
79° W
78° W
77° W
76° W
75° W
N
0
25
50
Miles
42° N
41° N
40° N
39° N
Cluster
11

Appendix C: Metrics and Index Calculation Examples

This appendix presents example metric calculations and proceeds step-by-step through the index development process using data from two samples: a sample from a 5th order site draining 84.5 square miles on Driftwood Branch Sinnemahoning Creek in Cameron County collected November 3, 2008; and a sample from a 1st order site draining 0.3 square miles in the headwaters of the West Branch Susquehanna River in Cambria County collected October 7, 2008. The taxa lists from the two sub-samples are below, followed by core metric and IBI calculations for each sample.

Driftwood Branch Sinnemahoning Creek
5th order @ 84.5 square miles
November 3, 2008

Taxa Name	Number of Individuals
Isonychia	15
Epeorus	10
Leucrocuta	13
Maccaffertium	18
Ephemerella	3
Eurylophella	3
Serratella	8
Paraleptophlebia	5
Stylogomphus	1
Taeniopteryx	13
Taenionema	37
Allocapnia	1
Neoperla	4
Paragnetina	2
Acroneuria	2
Nigronia	1
Chimarra	3
Polycentropus	2
Ceratopsyche	3
Cheumatopsyche	1
Rhyacophila	1
Glossosoma	2
Lepidostoma	2
Apatania	5
Neophylax	1
Oligochaeta	2
Psephenus	5
Optioservus	20
Atherix	2
Antocha	1
Chironomidae	6

West Branch Susquehanna River
1st order @ 0.3 square miles
October 7, 2008

Taxa Name	Number of Individuals
Baetis	5
Sweltsa	5
Sialis	1
Diplectrona	76
Rhyacophila	10
Oligochaeta	9
Optioservus	1
Chelifera	2
Tipula	5
Hexatoma	2
Limnophila	1
Prosimulium	3
Simulium	15
Chironomidae	68
Cambarus	2

Total Taxa Richness

Driftwood Branch Sinnemahoning Creek

= total number of taxa in a sub-sample

There are 31 taxa in this sub-sample.

Total Taxa Richness = 31

	Taxa Name	Number of Individuals
1	Isonychia	15
2	Epeorus	10
3	Leucrocuta	13
4	Maccaffertium	18
5	Ephemerella	3
6	Eurylophella	3
7	Serratella	8
8	Paraleptophlebia	5
9	Stylogomphus	1
10	Taeniopteryx	13
11	Taenionema	37
12	Allocapnia	1
13	Neoperla	4
14	Paragnetina	2
15	Acroneuria	2
16	Nigronia	1
17	Chimarra	3
18	Polycentropus	2
19	Ceratopsyche	3
20	Cheumatopsyche	1
21	Rhyacophila	1
22	Glossosoma	2
23	Lepidostoma	2
24	Apatania	5
25	Neophylax	1
26	Oligochaeta	2
27	Psephenus	5
28	Optioservus	20
29	Atherix	2
30	Antocha	1
31	Chironomidae	6

Total Taxa Richness

West Branch Susquehanna River

= total number of taxa in a sub-sample

There are 15 taxa in this sub-sample.

Total Taxa Richness = 15

	Taxa Name	Number of Individuals
1	Baetis	5
2	Sweltsa	5
3	Sialis	1
4	Diplectrona	76
5	Rhyacophila	10
6	Oligochaeta	9
7	Optioservus	1
8	Chelifera	2
9	Tipula	5
10	Hexatoma	2
11	Limnophila	1
12	Prosimulium	3
13	Simulium	15
14	Chironomidae	68
15	Cambarus	2

EPT Taxa Richness (PTV 0-4)

Driftwood Branch Sinnemahoning Creek

= number of taxa belonging to the insect orders Ephemeroptera, Plecoptera, or Trichoptera with **pollution tolerance values ≤ 4** in a sub-sample

There are **8** Ephemeroptera taxa **(PTV ≤ 4)** in this sub-sample.

Isoychia **Epeorus** **Leucrocuta**
Maccaffertium **Ephemerella** **Eurylophella**
Serratella **Paraleptophlebia**

There are **6** Plecoptera taxa **(PTV ≤ 4)** in this sub-sample.

Taeniopteryx **Taenionema** **Allocapnia**
Neoperla **Paragnetina** **Acroneuria**

There are **6** Trichoptera taxa **(PTV ≤ 4)** in this sub-sample.

Chimarra **Rhyacophila** **Glossosoma**
Lepidostoma **Apatania** **Neophylax**

EPT Taxa Richness (PTV 0-4) = 8 + 6 + 6
EPT Taxa Richness (PTV 0-4) **= 20**

Taxa Name	Number of Individuals	Pollution Tolerance Value
Isonychia	**15**	**3**
Epeorus	**10**	**0**
Leucrocuta	**13**	**1**
Maccaffertium	**18**	**3**
Ephemerella	**3**	**1**
Eurylophella	**3**	**4**
Serratella	**8**	**2**
Paraleptophlebia	**5**	**1**
Stylogomphus	1	4
Taeniopteryx	**13**	**2**
Taenionema	**37**	**3**
Allocapnia	**1**	**3**
Neoperla	**4**	**3**
Paragnetina	**2**	**1**
Acroneuria	**2**	**0**
Nigronia	1	2
Chimarra	**3**	**4**
Polycentropus	2	6
Ceratopsyche	3	5
Cheumatopsyche	1	6
Rhyacophila	**1**	**1**
Glossosoma	**2**	**0**
Lepidostoma	**2**	**1**
Apatania	**5**	**3**
Neophylax	**1**	**3**
Oligochaeta	2	10
Psephenus	5	4
Optioservus	20	4
Atherix	2	2
Antocha	1	3
Chironomidae	6	6

EPT Taxa Richness (PTV 0-4)

West Branch Susquehanna River

= number of taxa belonging to the insect orders Ephemeroptera, Plecoptera, or Trichoptera with pollution tolerance values ≤ 4 in a sub-sample

There are 0 Ephemeroptera taxa (**PTV ≤ 4**) in this sub-sample.

There is 1 Plecoptera taxa (**PTV ≤ 4**) in this sub-sample.
Sweltsa

There are 2 Trichoptera taxa (PTV ≤ 4) in this sub-sample.
Diplectrona **Rhyacophila**

EPT Taxa Richness (PTV 0-4) = 0 + 1 + 2
EPT Taxa Richness (PTV 0-4) **= 3**

Taxa Name	Number of Individuals	Pollution Tolerance Value
Baetis	5	6
Sweltsa	**5**	**0**
Sialis	1	6
Diplectrona	**76**	**0**
Rhyacophila	**10**	**1**
Oligochaeta	9	10
Optioservus	1	4
Chelifera	2	6
Tipula	5	4
Hexatoma	2	2
Limnophila	1	3
Prosimulium	3	2
Simulium	15	6
Chironomidae	68	6
Cambarus	2	6

Beck's Index (version 3)

Driftwood Branch Sinnemahoning creek

$$= 3(n_{taxaHILS0}) + 2(n_{taxaHILS1}) + 1(n_{taxaHILS2})$$

where $n_{taxaHILSi}$ = the number of taxa in a sub-sample with a pollution tolerance value (PTV) of i

There are **3 taxa** in this sub-sample with PTV = 0.
Epeorus **Acroneuria** **Glossosoma**

There are **6 taxa** in this sub-sample with PTV = 1.
Leucrocuta **Ephemerella** **Paraleptophlebia**
Paragnetina **Rhyacophila** **Lepidostoma**

There are **4 taxa** in this sub-sample with PTV = 2.
Serratella **Taeniopteryx** **Nigronia**
Atherix

Beck's Index (version 3) = 3(3) + 2(6) + 1(4)
Beck's Index (version 3) = 9 + 12 + 4
Beck's Index (version 3) **= 25**

Taxa Name	Number of Individuals	Pollution Tolerance Value
Isonychia	15	3
Epeorus	**10**	**0**
Leucrocuta	**13**	**1**
Maccaffertium	18	3
Ephemerella	**3**	**1**
Eurylophella	3	4
Serratella	**8**	**2**
Paraleptophlebia	**5**	**1**
Stylogomphus	1	4
Taeniopteryx	**13**	**2**
Taenionema	37	3
Allocapnia	1	3
Neoperla	4	3
Paragnetina	**2**	**1**
Acroneuria	**2**	**0**
Nigronia	**1**	**2**
Chimarra	3	4
Polycentropus	2	6
Ceratopsyche	3	5
Cheumatopsyche	1	6
Rhyacophila	**1**	**1**
Glossosoma	**2**	**0**
Lepidostoma	**2**	**1**
Apatania	5	3
Neophylax	1	3
Oligochaeta	2	10
Psephenus	5	4
Optioservus	20	4
Atherix	**2**	**2**
Antocha	1	3
Chironomidae	6	6

Beck's Index (version 3)

West Branch Susquehanna River

$= 3(n_{taxaHILS0}) + 2(n_{taxaHILS1}) + 1(n_{taxaHILS2})$

where $n_{taxaHILSi}$ = the number of taxa in a sub-sample with a pollution tolerance value (PTV) of i

There are **2 taxa** in this sub-sample with PTV = 0.
Sweltsa **Diplectrona**

There is **1 taxa** in this sub-sample with PTV = 1.
Rhyacophila

There are **2 taxa** in this sub-sample with PTV = 2.
Hexatoma **Prosimulium**

Beck's Index (version 3) = 3(2) + 2(1) + 1(2)
Beck's Index (version 3) = 6 + 2 + 2
Beck's Index (version 3) **= 10**

Taxa Name	Number of Individuals	Pollution Tolerance Value
Baetis	5	6
Sweltsa	**5**	**0**
Sialis	1	6
Diplectrona	**76**	**0**
Rhyacophila	**10**	**1**
Oligochaeta	9	10
Optioservus	1	4
Chelifera	2	6
Tipula	5	4
Hexatoma	**2**	**2**
Limnophila	1	3
Prosimulium	**3**	**2**
Simulium	15	6
Chironomidae	68	6
Cambarus	2	6

Hilsenhoff Biotic Index

Driftwood Branch Sinnemahoning Creek

$$= \sum_{i=0}^{10} [(i * n_{indvPTVi})] / N$$

where $n_{indvPTVi}$ = the number of individuals in a sub-sample with pollution tolerance value (PTV) of i and N = the total number of individuals in a sub-sample

There are 14 individuals with PTV = 0
There are 26 individuals with PTV = 1
There are 24 individuals with PTV = 2
There are 82 individuals with PTV = 3
There are 32 individuals with PTV = 4
There are 3 individuals with PTV = 5
There are 9 individuals with PTV = 6
There are 0 individuals with PTV = 7, 8, or 9
There are 2 individuals with PTV = 10.

There are a total of 192 individuals in the sub-sample.

Hilsenhoff Biotic Index =

[(0 * 14) + (1 * 26) + (2 * 24) +
(3 * 82) + (4 * 32) + (5 * 3) +
(6 * 9) + (7 * 0) + (8 * 0) +
(9 * 0) + (10 * 2)] / 192

Hilsenhoff Biotic Index = 2.80

Taxa Name	Number of Individuals	Pollution Tolerance Value
Isonychia	15	3
Epeorus	10	0
Leucrocuta	13	1
Maccaffertium	18	3
Ephemerella	3	1
Eurylophella	3	4
Serratella	8	2
Paraleptophlebia	5	1
Stylogomphus	1	4
Taeniopteryx	13	2
Taenionema	37	3
Allocapnia	1	3
Neoperla	4	3
Paragnetina	2	1
Acroneuria	2	0
Nigronia	1	2
Chimarra	3	4
Polycentropus	2	6
Ceratopsyche	3	5
Cheumatopsyche	1	6
Rhyacophila	1	1
Glossosoma	2	0
Lepidostoma	2	1
Apatania	5	3
Neophylax	1	3
Oligochaeta	2	10
Psephenus	5	4
Optioservus	20	4
Atherix	2	2
Antocha	1	3
Chironomidae	6	6

Hilsenhoff Biotic Index

West Branch Susquehanna River

$$= \sum_{i=0}^{10} [(i * n_{indvPTVi})] / N$$

where $n_{indvPTVi}$ = the number of individuals in a sub-sample with pollution tolerance value (PTV) of i and N = the total number of individuals in a sub-sample

There are 81 individuals with PTV = 0
There are 10 individuals with PTV = 1
There are 5 individuals with PTV = 2
There is 1 individual with PTV = 3
There are 6 individuals with PTV = 4
There are 0 individuals with PTV = 5
There are 93 individuals with PTV = 6
There are 0 individuals with PTV = 7, 8, or 9
There are 9 individuals with PTV = 10.

There are a total of 205 individuals in the sub-sample.

Hilsenhoff Biotic Index =

[(0 * 81) + (1 * 10) + (2 * 5) +
(3 * 1) + (4 * 6) + (5 * 0) +
(6 * 93) + (7 * 0) + (8 * 0) +
(9 * 0) + (10 * 9)] / 205

Hilsenhoff Biotic Index = 3.39

Taxa Name	Number of Individuals	Pollution Tolerance Value
Baetis	5	6
Sweltsa	5	0
Sialis	1	6
Diplectrona	76	0
Rhyacophila	10	1
Oligochaeta	9	10
Optioservus	1	4
Chelifera	2	6
Tipula	5	4
Hexatoma	2	2
Limnophila	1	3
Prosimulium	3	2
Simulium	15	6
Chironomidae	68	6
Cambarus	2	6

Shannon Diversity

Driftwood Branch Sinnemahoning Creek

$$= [- \sum_{i=1}^{Rich} (n_i / N) \ln (n_i / N)]$$

where n_i = the number of individuals in each taxon (relative abundance); N = the total number of individuals in a sub-sample; and Rich = the total number of taxa in a sub-sample (total taxa richness)

There are 31 taxa in this sub-sample. The numbers of individuals in each taxon are listed in the table to the right. There are a total of 192 individuals in the sub-sample. Starting at the top of the taxa list and working down – row by row, taxon by taxon – this metric is calculated as,

Shannon Diversity =

– (15 / 192) ln (15 / 192) +
(10 / 192) ln (10 / 192) +
(13 / 192) ln (13 / 192) +
(18 / 192) ln (18 / 192) +
(3 / 192) ln (3 / 192) +
(3 / 192) ln (3 / 192) +
(8 / 192) ln (8 / 192) +
... *(do this for all 31 taxa)* ...
(6 / 192) ln (6 / 192)

Shannon Diversity = 2.88

Taxa Name	Number of Individuals
Isonychia	15
Epeorus	10
Leucrocuta	13
Maccaffertium	18
Ephemerella	3
Eurylophella	3
Serratella	8
Paraleptophlebia	5
Stylogomphus	1
Taeniopteryx	13
Taenionema	37
Allocapnia	1
Neoperla	4
Paragnetina	2
Acroneuria	2
Nigronia	1
Chimarra	3
Polycentropus	2
Ceratopsyche	3
Cheumatopsyche	1
Rhyacophila	1
Glossosoma	2
Lepidostoma	2
Apatania	5
Neophylax	1
Oligochaeta	2
Psephenus	5
Optioservus	20
Atherix	2
Antocha	1
Chironomidae	6

Shannon Diversity

West Branch Susquehanna River

$$= [- \sum_{i=1}^{Rich} (n_i / N) \ln (n_i / N)]$$

where n_i = the number of individuals in each taxon (relative abundance); N = the total number of individuals in a sub-sample; and Rich = the total number of taxa in a sub-sample (total taxa richness)

There are 15 taxa in this sub-sample. The numbers of individuals in each taxon are listed in the table to the right. There are a total of 205 individuals in the sub-sample. Starting at the top of the taxa list and working down – row by row, taxon by taxon – this metric is calculated as,

Taxa Name	Number of Individuals
Baetis	5
Sweltsa	5
Sialis	1
Diplectrona	76
Rhyacophila	10
Oligochaeta	9
Optioservus	1
Chelifera	2
Tipula	5
Hexatoma	2
Limnophila	1
Prosimulium	3
Simulium	15
Chironomidae	68
Cambarus	2

Shannon Diversity =

– (5 / 205) ln (5 / 205) +
(5 / 205) ln (5 / 205) +
(1 / 205) ln (1 / 205) +
(76 / 205) ln (76 / 205) +
(10 / 205) ln (10 / 205) +
(9 / 205) ln (9 / 205) +
(1 / 205) ln (1 / 205) +
... *(do this for all 13 taxa)* ...
(2 / 205) ln (2 / 205)

Shannon Diversity = 1.76

Percent Sensitive Individuals (PTV 0-3 only)

Driftwood Branch Sinnemahoning Creek

$$= \left(\sum_{i=0}^{3} n_{indvPTVi} \right) / N * 100$$

where $n_{indvPTVi}$ = the number of individuals in a sub-sample with pollution tolerance value (PTV) of i and N = the total number of individuals in a sub-sample

There are **14 individuals with PTV = 0**
There are **26 individuals with PTV = 1**
There are **24 individuals with PTV = 2**
There are **82 individuals with PTV = 3**

There are a **total of 192 individuals in the sub-sample.**

Percent Sensitive Individuals (PTV 0-3 only) =
(**14** + **26** + **24** + **82**) / **192** *100

Percent Sensitive Individuals (PTV 0-3 only) =
146 / 192 * 100

Percent Sensitive Individuals (PTV 0-3 only) **= 76.0%**

Taxa Name	Number of Individuals	Pollution Tolerance Value
Isonychia	**15**	**3**
Epeorus	**10**	**0**
Leucrocuta	**13**	**1**
Maccaffertium	**18**	**3**
Ephemerella	**3**	**1**
Eurylophella	3	4
Serratella	**8**	**2**
Paraleptophlebia	**5**	**1**
Stylogomphus	1	4
Taeniopteryx	**13**	**2**
Taenionema	**37**	**3**
Allocapnia	**1**	**3**
Neoperla	**4**	**3**
Paragnetina	**2**	**1**
Acroneuria	**2**	**0**
Nigronia	**1**	**2**
Chimarra	3	4
Polycentropus	2	6
Ceratopsyche	3	5
Cheumatopsyche	1	6
Rhyacophila	**1**	**1**
Glossosoma	**2**	**0**
Lepidostoma	**2**	**1**
Apatania	**5**	**3**
Neophylax	**1**	**3**
Oligochaeta	2	10
Psephenus	5	4
Optioservus	20	4
Atherix	**2**	**2**
Antocha	**1**	**3**
Chironomidae	6	6

Percent Sensitive Individuals (PTV 0-3 only)

West Branch Susquehanna River

$$= \left(\sum_{i=0}^{3} n_{indvPTVi} \right) / N * 100$$

where $n_{indvPTVi}$ = the number of individuals in a sub-sample with pollution tolerance value (PTV) of i and N = the total number of individuals in a sub-sample

There are **81 individuals with PTV = 0**
There are **10 individuals with PTV = 1**
There are **5 individuals with PTV = 2**
There is **1 individual with PTV = 3**

There are a **total of 205 individuals in the sub-sample.**

Percent Sensitive Individuals (PTV 0-3 only) =
(**81** + **10** + **5** + **1**) / **205** *100

Percent Sensitive Individuals (PTV 0-3 only) =
97 / 205 * 100

Percent Sensitive Individuals (PTV 0-3 only) **= 47.3%**

Taxa Name	Number of Individuals	Pollution Tolerance Value
Baetis	5	6
Sweltsa	**5**	**0**
Sialis	1	6
Diplectrona	**76**	**0**
Rhyacophila	**10**	**1**
Oligochaeta	9	10
Optioservus	1	4
Chelifera	2	6
Tipula	5	4
Hexatoma	**2**	**2**
Limnophila	**1**	**3**
Prosimulium	**3**	**2**
Simulium	15	6
Chironomidae	68	6
Cambarus	2	6

Metric Standardization and Index Scoring

Table D1 lists the small-stream and large-stream standardization values for each core metric.

Table D1. Values used to standardize core metrics

Metric	Metric Standardization Values	
	Smaller streams (1^{st} to 3^{rd} order, < 25 square miles)	**Larger streams** (5^{th} order and larger, > 50 square miles)
Total Taxa Richness	33	31
EPT Taxa Richness (PTV 0-4 only)	19	16
Beck's Index (version 3)	38	22
Hilsenhoff Biotic Index	1.89	3.05
Shannon Diversity	2.86	2.86
% Sensitive Individuals (PTV 0-3 only)	84.5	66.7

The Hilsenhoff Biotic Index metric values are expected to increase in value with increasing anthropogenic stress and are standardized using the following equation:

Hilsenhoff Biotic Index standardized score = (10 – observed value) / (10 – standardization value) * 100

The other five core metrics values are expected to decrease in value with increasing anthropogenic stress and are standardized using the following equation:

Standardized metric score = observed value / standardization value * 100

Table D2 and Table D3 show the standardization and index scoring calculations for the two samples discussed above.

Table D2. Standardization and index calculations for the Driftwood Branch Sinnemahoning Creek sample. The large-stream standardization values are used here because the sample is from a 5th order site draining 84.5 square miles of land.

Metric	Standardization Equation	Observed Metric Value	Standardized Metric Score	Adjusted Standardized Metric Score Maximum = 100
Total Taxa Richness	observed value / **31** * 100	31	100.0	100
EPT Taxa Richness (PTV 0-4 only)	observed value / **16** * 100	20	125.0	100
Beck's Index (version 3)	observed value / **22** * 100	25	113.6	100
Hilsenhoff Biotic Index	(10 – observed value) / (10 – **3.05**) * 100	2.80	103.6	100
Shannon Diversity	observed value / **2.86** * 100	2.88	100.7	100
Percent Sensitive Individuals (PTV 0-3 only)	observed value / **66.7** * 100	76.0	113.9	100
Arithmetic average of adjusted standardized core metric scores = **IBI Score =**				**100.0**

Table D3. Standardization and index calculations for the West Branch Susquehanna River sample. The small-stream standardization values are used here because the sample is from a 1st order site draining 0.3 square miles of land.

Metric	Standardization Equation	Observed Metric Value	Standardized Metric Score	Adjusted Standardized Metric Score Maximum = 100
Total Taxa Richness	observed value / **33** * 100	15	45.5	45.5
EPT Taxa Richness (PTV 0-4 only)	observed value / **19** * 100	3	15.8	15.8
Beck's Index (version 3)	observed value / **38** * 100	10	26.3	26.3
Hilsenhoff Biotic Index	(10 – observed value) / (10 – **1.89**) * 100	3.39	81.5	81.5
Shannon Diversity	observed value / **2.86** * 100	1.76	61.5	61.5
Percent Sensitive Individuals (PTV 0-3 only)	observed value / **84.5** * 100	47.3	56.0	56.0
Arithmetic average of adjusted standardized core metric scores = **IBI Score =**				**47.8**

Appendix D: Table of Taxa

The following table lists the pollution tolerance value (PTV), BCG attributes, and functional feeding group (FFG) assignment used by PADEP for each benthic macroinvertebrate taxon. The FFG abbreviations stand for collector-gatherer (CG), filter-collector (FC), piercer (PI), predator (PR), scraper (SC), shredder (SH), and unknown (UK). Note that some taxa were assigned different BCG attributes for smaller streams and for larger streams.

Taxa	PTV	BCG attribute		FFG
		small-stream	large-stream	
Insecta				
Collembola	**9**			**CG**
Onychiuridae	**9**			**CG**
Onychiurus	9			CG
Poduridae	**9**			**CG**
Podura	9			CG
Ephemeroptera				
Ameletidae	**0**			**CG**
Ameletus	0	2	2	CG
Siphlonurus	7			CG
Metrotopus	2			CG
Siphloplecton	2	2	2	CG
Baetidae	**6**	**3**	**3**	**CG**
Acentrella	4	3	3	SC
Acerpenna	6	3	3	CG
Baetis	6	4	5	CG
Barbaetis	6			CG
Callibaetis	9	4	4	CG
Centroptilum	2	3	3	CG
Cloeon	4	3	3	CG
Diphetor	6	2	2	CG
Fallceon	6			CG
Procloeon	6	4	4	CG
Heterocloeon	2	3	3	SC
Plauditus	4			CG
Pseudocloeon	4	3	3	CG
Isonychiidae	**3**			**CG**
Isonychia	3	3	3	CG
Heptageniidae	**3**			**SC**
Epeorus	0	2	2	SC
Heptagenia	4	2	3	SC
Leucrocuta	1	3	3	SC
Nixe	2	1	1	SC
Rhithrogena	0	2	2	CG
Stenacron	4	4	4	SC
Stenonema(old genus)	3	3	3	SC
Stenonema	4	4	4	SC
Maccaffertium	3	3	3	SC
Cinygmula	1	1	1	CG
Arthropleidae	**3**			**SC**
Arthroplea	3			SC
Ephemerellidae	**2**			**CG**
Attenella	2	2	2	SC
Drunella	1	2	2	SC
Ephemerella	1	3	2	CG
Eurylophella	4	3	2	SC
Serratella	2	3	3	CG
Dannella	3	3	3	CG
Neoephemeridae	**3**			**CG**
Neoephemera	3			CG
Caenidae	**7**			**CG**
Brachycercus	3			CG
Caenis	7	5	5	CG
Baetiscidae	**3**			**CG**
Baetisca	4	2	2	CG
Leptophlebiidae	**4**	**2**	**2**	**CG**
Choroterpes	2	2	2	CG
Habrophlebia	4	3	3	CG
Habrophlebiodes	6	2	2	SC
Leptophlebia	4	3	3	CG
Paraleptophlebia	1	2	2	CG
Anthopotamus	4	3	3	CG
Ephemeridae	**4**			**CG**
Ephemera	2	3	2	CG
Hexagenia	6	4	4	CG
Litobrancha	6	1	1	CG

Pentagenia	4			CG
Polymitarcyidae	**2**			**CG**
Ephoron	2	3	3	CG
Tricorythidae	**4**			**CG**
Tricorythodes	4	5	5	CG
Leptohyphes	4			CG
Odonata				**PR**
Petaluridae	**5**			**PR**
Tachopteryx	5			PR
Gomphidae	**4**	**3**	**3**	**PR**
Aphylla	4			PR
Arigomphus	4	4	4	PR
Dromogomphus	4	4	4	PR
Gomphus	5	4	4	PR
Hagenius	3	3	3	PR
Lanthus	5	2	2	PR
Ophiogomphus	1	3	3	PR
Progomphus	5	3	3	PR
Stylogomphus	4	4	4	PR
Stylurus	4	4	4	PR
Aeshnidae	**3**			**PR**
Aeshna	5	4	4	PR
Anax	5	4	4	PR
Basiaeschna	2	4	4	PR
Boyeria	2	3	3	PR
Epiaeschna	2			PR
Gomphaeschna	2	4	4	PR
Nasiaeschna	2			PR
Cordulegastridae	**3**			**PR**
Cordulegaster	3	3	3	PR
Corduliidae	**5**			**PR**
Didymops	4	4	4	PR
Cordulia	4			PR
Dorocordulia	4			PR
Epitheca	4			PR
Helocordulia	2			PR
Somatochlora	1	2	2	PR
Williamsonia	4			PR
Macromia	2	4	4	PR
Neurocordulia	3			PR
Libellulidae	**9**			**PR**
Celithemis	2			PR
Erythemis	5			PR

Erythrodiplax	5			PR
Ladona	6			PR
Leucorrhinia	6			PR
Libellula	8			PR
Nannothemis	6			PR
Pachydiplax	8			PR
Pantala	7			PR
Perithemis	4			PR
Plathemis	3			PR
Sympetrum	4			PR
Tramea	4			PR
Calopterygidae	**5**	**4**	**4**	**PR**
Calopteryx	6	4	4	PR
Hetaerina	6	4	4	PR
Lestes	9			PR
Coenagrionidae	**8**	**4**	**4**	**PR**
Amphiagrion	5			PR
Argia	6	4	4	PR
Chromagrion	4			PR
Enallagma	8	4	4	PR
Ischnura	9	4	4	PR
Nehalennia	7			PR
Plecoptera				**PR**
Pteronarcyidae	**0**			**SH**
Pteronarcys	0	1	2	SH
Peltoperlidae	**2**	**2**	**2**	**SH**
Peltoperla	2	1	1	SH
Tallaperla	0	1	1	SH
Viehoperla	2			SH
Taeniopterygidae	**2**	**3**	**3**	**SH**
Taeniopteryx	2	3	3	SH
Bolotoperla	2			SH
Oemopteryx	3	2	2	SH
Strophopteryx	3	3	3	SH
Taenionema	3	1	1	SH
Nemouridae	**2**	**3**	**3**	**SH**
Amphinemura	3	3	3	SH
Ostrocerca	2	1	1	SH
Paranemoura	2			SH
Podmosta	2			SH
Prostoia	2	3	3	SH
Shipsa	2	1	1	SH
Soyedina	0	1	1	SH

Zapada	2			SH
Nemoura	1	1	1	SH
Leuctridae	**0**	**3**	**3**	**SH**
Megaleuctra	0			SH
Leuctra	0	2	2	SH
Paraleuctra	0	1	1	SH
Zealeuctra	0			SH
Capniidae	**3**	**3**	**3**	**SH**
Allocapnia	3	3	3	SH
Capnia	1			SH
Nemocapnia	1			SH
Paracapnia	1	2	2	SH
Utacapnia	1			SH
Capnura	1			SH
Perlidae	**3**	**3**	**3**	**PR**
Agnetina	2	3	3	PR
Hansonoperla	3			PR
Neoperla	3	2	2	PR
Paragnetina	1	2	2	PR
Acroneuria	0	3	3	PR
Attaneuria	3	2	2	PR
Eccoptura	2	2	2	PR
Perlesta	4	3	3	PR
Perlinella	2	2	2	PR
Perlodidae	**2**	**2**	**2**	**PR**
Cultus	2	1	1	PR
Diploperla	2	2	2	PR
Diura	2			PR
Helopicus	2	3	3	PR
Hydroperla	1			PR
Isogenoides	0	1	1	PR
Malirekus	2	1	1	PR
Oconoperla	2			PR
Remenus	2	1	1	PR
Yugus	2	1	1	PR
Clioperla	2			PR
Isoperla	2	2	2	PR
Arcynopteryx	2			PR
Chloroperlidae	**0**	**2**	**2**	**PR**
Utaperla	0			PR
Alloperla	0	1	1	CG
Haploperla	0	3	3	PR
Rasvena	0	1	1	PR

Suwallia	0	1	1	CG
Sweltsa	0	3	3	PR
Hemiptera				
Hydrometridae	**9**			**PR**
Veliidae	**8**			**PR**
Microvelia	9			PR
Rhagovelia	9			PR
Steinovelia	9			PR
Ceratocombidae	**9**			**PR**
Ceratocombus	9			PR
Gerridae	**9**			**PR**
Aquarius	9			PR
Gerris	9			PR
Halobates	9			PR
Rheumatobates	9			PR
Metrobates	9			PR
Trepobates	9			PR
Limnoporus	9			PR
Belostomatidae	**9**			**PR**
Belostoma	9			PR
Lethocerus	9			PR
Nepidae	**8**			**PR**
Nepa	8			PR
Ranatra	8			PR
Pleidae	**8**			**PR**
Neoplea	8			PR
Naucoridae	**8**			**PR**
Pelocoris	8			PR
Corixidae	**8**	**5**	**5**	**PR**
Hesperocorixa	5	5	5	PR
Palmacorixa	8	4	4	PR
Ramphocorixa	8	4	4	PR
Sigara	8	4	4	PR
Trichocorixa	8	5	5	PR
Notonectidae	**8**			**PR**
Buenoa	8			PR
Notonecta	8			PR
Mesoveliidae	**9**			**PR**
Mesovelia	9			PR
Hebridae	**8**			**PR**
Hebrus	8			PR
Merragata	8			PR
Saldidae	**8**			**PR**

Micracanthia	8			PR
Pentacora	8			PR
Salda	8			PR
Saldula	8			PR
Gelastocoridae	**8**			**PR**
Gelastocoris	8			PR
Ochteridae	**8**			**PR**
Ochterus	8			PR
Megaloptera	**8**			**PR**
Sialis	6	5	5	PR
Corydalidae	**3**			**PR**
Corydalus	4	4	4	PR
Chauliodes	4	4	4	PR
Neohermes	2			PR
Nigronia	2	3	3	PR
Neuroptera	**3**			**PR**
Sisyridae	**1**			**PI**
Climacia	1			PI
Sisyra	1			PI
Trichoptera				
Philopotamidae	**3**			**FC**
Chimarra	4	4	4	FC
Dolophilodes	0	2	2	FC
Wormaldia	0	1	1	FC
Psychomyiidae	**2**	**3**	**3**	**CG**
Lype	2	2	2	CG
Psychomyia	2	3	3	CG
Polycentropodidae	**6**			**FC**
Cernotina	6			PR
Cyrnellus	8	5	5	FC
Neureclipsis	7	3	3	FC
Polycentropus	6	4	4	FC
Phylocentropus	5	4	4	FC
Nyctiophylax	5	4	4	PR
Hydropsychidae	**5**			**FC**
Arctopsyche	1			FC
Parapsyche	0	1	1	FC
Diplectrona	0	2	2	FC
Homoplectra	5			FC
Ceratopsyche	5	4	4	FC
Cheumatopsyche	6	5	5	FC
Hydropsyche	5	5	5	FC
Potamyia	5	3	3	FC

Macrostemum	3	4	4	FC
Rhyacophilidae	**1**			**SC**
Rhyacophila	1	2	2	PR
Glossosomatidae	**0**	**3**	**3**	**SC**
Glossosoma	0	3	3	SC
Agapetus	0	3	3	SC
Culoptila	1	3	3	SC
Protoptila	1	2	2	SC
Hydroptilidae	**4**			**PI**
Palaeagapetus	1	1	1	SH
Agraylea	8	4	4	CG
Dibusa	4			SC
Hydroptila	6	5	5	SC
Ochrotrichia	4			SC
Oxyethira	3	2	2	CG
Stactobiella	2			SC
Leucotrichia	6	4	4	SC
Ithytrichia	6			SC
Orthotrichia	6			SH
Neotrichia	2			SC
Mayatrichia	4			SC
Phryganeidae	**4**			**SH**
Agrypnia	7			SH
Banksiola	2			SH
Fabria	4			SH
Hagenella	5			SH
Oligostomis	5			SH
Phryganea	8			SH
Ptilostomis	5	2	2	SH
Brachycentridae	**1**	**2**	**2**	**FC**
Adicrophleps	2	1	1	SH
Brachycentrus	1	3	3	FC
Micrasema	2	3	3	SH
Lepidostomatidae	**1**	**2**	**2**	**SH**
Lepidostoma	1	2	2	SH
Limnephilidae	**4**	**3**	**3**	**SH**
Ironoquia	3			SH
Onocosmoecus	3			SH
Apatania	3	2	2	SC
Pseudostenophylax	0	3	3	SH
Anabolia	5			SH
Arctopora	5			SH
Clostoeca	5			SH

Frenesia	4			SH
Hesperophylax	4	3	3	CG
Hydatophylax	2	2	2	SH
Leptophylax	2			SH
Limnephilus	3	3	3	SH
Philarctus	3			SH
Platycentropus	4	3	3	SH
Pycnopsyche	4	3	3	SH
Goera	0	1	1	SC
Madeophylax	4			SH
Glyphopsyche	3			SH
Uenoidae	**3**			**SC**
Neophylax	3	3	3	SC
Beraeidae	**3**			**SC**
Beraea	3			SC
Sericostomatidae	**3**			**SH**
Agarodes	3	2	2	SH
Psilotreta	0	1	1	SC
Molannidae	**6**			**SC**
Molanna	6	2	2	SC
Helicopsychidae	**3**			**SC**
Helicopsyche	3	3	3	SC
Calamoceratidae	**5**			**SH**
Heteroplectron	5	1	1	SH
Leptoceridae	**4**			**PR**
Ceraclea	3	3	3	CG
Leptocerus	3			SH
Mystacides	4	3	3	CG
Nectopsyche	3	3	3	SH
Oecetis	8	3	3	PR
Setodes	2	2	2	CG
Triaenodes	6	3	3	SH
Odontoceridae	**0**	**1**	**1**	**SH**
Lepidoptera	**5**			**SH**
Pyralidae	**5**			**SH**
Langessa	5			SH
Munroessa	5			SH
Neocataclysta	5			SH
Nymphula	7			SH
Nymphuliella	5			SH
Parapoynx	5			SH
Synclita	5			FC
Eoparargyractis	5			SH

Petrophila	5	5	5	SC
Acentria	5			SH
Schoenobius	5			SH
Chilo	5			SH
Acigona	5			SH
Ostrinia	5			SH
Nepticulidae	**5**			**SH**
Stigmella	5			SH
Cosmopterigidae	**5**			**SH**
Cosmopteryx	5			SH
Lymnaecia	5			SH
Noctuidae	**5**			**SH**
Archanara	5			SH
Bellura	5			SH
Simyra	5			SH
Tortricidae	**5**			**SH**
Archips	5			SH
Coleophoridae	**6**			**SH**
Colephora	6			SH
Coleoptera				
Gyrinidae	**4**	**4**	**4**	**PR**
Dineutus	4	4	4	PR
Gyrinus	4	4	4	PR
Spanglerogyrus	4			PR
Haliplidae	**5**			**SH**
Haliplus	5			SH
Peltodytes	5			SH
Dytiscidae	**5**	**4**	**4**	**PR**
Acilius	5	4	4	PR
Agabetes	5	4	4	PR
Agabus	5	4	4	PR
Bidessonotus	5			PR
Brachyvatus	5			PR
Celina	5			PR
Copelatus	5	4	4	PR
Colymbetes	5			PR
Coptotomus	5			PR
Cybister	5	4	4	PR
Desmopachria	5			PR
Dytiscus	5			PR
Graphoderus	5			PR
Hydaticus	5			PR
Hydrovatus	5	4	4	PR

Hygrotus	5			PR
Ilybius	5	4	4	PR
Laccophilus	5	4	4	PR
Laccornis	5			PR
Liodessus	5			PR
Lioporius	5			PR
Matus	5			PR
Nebrioporus	5			PR
Oreodytes	5			PR
Rhantus	5			PR
Stictotarsus	5			PR
Uvarus	5	4	4	PR
Hydroporus	5	4	4	PR
Noteridae	**5**			**PR**
Hydrocanthus	5			PR
Pronoterus	5			PR
Suphis	5			PR
Suphisellus	5			PR
Helophoridae	**5**			**SH**
Helophorus	5			SH
Hydrochidae	**5**			**SH**
Hydrochus	5			SH
Hydrophilidae	**5**			**PR**
Anacaena	5			PR
Berosus	5	5	5	PR
Chaetarthria	5			PR
Crenitis	5			PR
Cymbiodyta	5			PR
Derallus	5			PR
Dibolocelus	*5*			*PR*
Enochrus	5			PR
Helochares	5			PR
Helocombus	5			PR
Hydrobius	5			PR
Hydrochara	5			PR
Hydrophilus	5			PR
Laccobius	5			PR
Paracymus	5			PR
Sperchopsis	5			PR
Tropisternus	5			PR
Staphylinidae	**5**			**PR**
Bledius	5			PR
Carpelimus	5			PR

Psephidonus	5			PR
Thinobius	5			PR
Stenus	5			PR
Hydraenidae	**6**			**CG**
Hydraena	6			CG
Limnebius	6			CG
Ochthebius	6			CG
Psephenidae	**4**			**SC**
Eubrianax	4			SC
Psephenus	4	4	4	SC
Dicranopselaphus	4			SC
Ectopria	5	3	3	SC
Dryopidae	**5**			**SC**
Dryops	5			SC
Helichus	5	4	4	SC
Scirtidae	**8**			**SC**
Cyphon	8			SC
Elodes	8			SC
Flavohelodes	8			SC
Scirtes	8			SC
Elmidae	**5**			**CG**
Ancyronyx	2	4	4	CG
Dubiraphia	6	4	4	SC
Gonielmis	5			SC
Macronychus	2	4	4	SC
Microcylloepus	2	4	4	SC
Optioservus	4	4	4	SC
Ordobrevia	5			SC
Oulimnius	5	3	2	SC
Promoresia	2	3	2	SC
Stenelmis	5	5	5	SC
Anchytarsus	5	3	2	SH
Lutrochidae	**6**			**UK**
Lutrochus	6			UK
Chrysomelidae	**5**			**SH**
Disonycha	5			SH
Donacia	5			SH
Hydrothassa	*5*			*SH*
Neohaemonia	5			SH
Prasocuris	5			SH
Pyrrhalta	5			SH
Curculionidae	**6**			**SH**
Auleutes	6			SH

Bagous	6			SH
Brachybamus	6			SH
Euhrychiopsis	6			SH
Lissorhoptrus	6			SH
Listronotus	6			SH
Lixellus	6			SH
Lixus	6			SH
Notiodes	6			SH
Onychylis	6			SH
Perenthis	6			SH
Pelenomus	6			SH
Phytobius	6			SH
Stenopelmus	6			SH
Steremnius	6			SH
Tanysphyrus	6			SH
Histeridae	**5**			**SH**
Pompilidae	**5**			**UK**
Anoplius	5			UK
Scelionidae	**5**			**UK**
Pseudanteris	5			UK
Telenomus	5			UK
Thoron	5			UK
Tiphodytes	5			UK
Diapriidae	**5**			**UK**
Trichopria	5			UK
Ichneumonidae	**5**			**UK**
Apsilops	5			UK
Cremastus	5			UK
Medophron	5			UK
Mesoleptus	5			UK
Phygadeuon	5			UK
Braconidae	**5**			**UK**
Ademon	5			UK
Aphanta	5			UK
Asobara	5			UK
Bracon	5			UK
Chaenusa	5			UK
Chorebidella	5			UK
Chorebus	5			UK
Dacnusa	5			UK
Opius	5			UK
Phaenocarpa	5			UK
Mymaridae	**5**			**UK**

Caraphractus	5			UK
Trichogrammatida	**5**			**UK**
Hydrophylita	5			UK
Lathromeroidea	5			UK
Paracentrobia	5			UK
Trichogramma	5			UK
Eulophidae	**5**			**UK**
Aprostocetus	5			UK
Mestocharis	5			UK
Tetrastichus	5			UK
Pteromalidae	**5**			**UK**
Gyrinophagus	5			UK
Sisridivora	5			UK
Eucoilidae	**5**			**UK**
Hexacola	5			UK
Diptera				
Blephariceridae	**0**			**SC**
Blepharicera	0	1	1	SC
Ceratopogonidae	**6**	**4**	**4**	**PR**
Dasyhelea	6	4	4	CG
Atrichopogon	2	4	4	PR
Forcipomyia	6	4	4	SC
Alluaudomyia	6	4	4	PR
Bezzia	6	4	4	PR
Brachypogon	6			PR
Ceratopogon	6	4	4	PR
Clinohelea	6			PR
Culicoides	10	4	4	PR
Johannsenomyia	6			PR
Mallochohelea	6	4	4	PR
Monohelea	6			PR
Nilobezzia	6			PR
Palpomyia	6	4	4	PR
Probezzia	6	4	4	PR
Serromyia	6			PR
Sphaeromias	6			PR
Stilobezzia	6	4	4	PR
Leptoconops	6			PR
Chaoboridae	**8**			**PR**
Chaoborus	8			PR
Mochlonyx	8			PR
Dixidae	**1**	**2**	**2**	**CG**
Dixa	1	2	2	CG

Dixella	1			CG
Nymphomyiidae	**6**			**SC**
Palaeodipteron	6			SC
Psychodidae	**10**	**5**	**5**	**CG**
Pericoma	4	5	5	CG
Philosepedon	10			CG
Psychoda	10	5	5	CG
Telmatoscopus	10	5	5	CG
Threticus	10			CG
Ptychopteridae	**8**			**CG**
Bittacomorpha	8			CG
Bittacomorphella	8			CG
Ptychoptera	8			CG
Protoplasa	6			CG
Thaumalea	6			SC
Trichothaumalea	6			SC
Athericidae	**2**			**PR**
Atherix	2	3	3	PR
Pelecorhynchidae	**5**			**PR**
Glutops	5			PR
Dolichopodidae	**4**			**PR**
Argyra	4			PR
Asyndetus	4			PR
Campsicnemus	4			CG
Dolichopus	4			PR
Hercostomus	4			PR
Hydrophorus	4			PR
Hypocharassus	4			PR
Liancalus	4			PR
Pelastoneurus	4			PR
Sympycnus	4			PR
Tachytrechus	4			PR
Telmaturgus	4			PR
Thinophilus	4			PR
Empididae	**6**	**4**	**4**	**PR**
Chelifera	6	4	4	PR
Chelipoda	6			PR
Clinocera	6	4	4	PR
Dolichocephala	5			PR
Hemerodromia	6	4	4	PR
Metachela	6			PR
Neoplasta	6			PR
Oreothalia	6			PR

Proclinopyga	6			PR
Rhamphomyia	6			PR
Roederiodes	6			PR
Stilpon	6			PR
Trichoclinocera	6			PR
Oreogeton	6			PR
Stratiomyidae	**8**	**6**	**6**	**CG**
Caloparyphus	8			CG
Euparyphus	8			CG
Hedriodiscus	8			SC
Labostigmina	8			CG
Nemotelus	8			CG
Odontomyia	8			CG
Oxycera	8			SC
Sargus	8			CG
Stratiomys	5			CG
Tabanidae	**6**	**5**	**5**	**PI**
Atylotus	6			PI
Chrysops	7	5	5	PI
Haematopota	6			PR
Hybomitra	6			PR
Merycomyia	6			PR
Tabanus	5	5	5	PR
Diachlorus	6			PR
Ephydridae	**6**	**5**	**5**	**PI**
Leptopsilopa	6			CG
Psilopa	6			CG
Rhysophora	6			SH
Muscidae	**6**			**PR**
Caricea	6			PR
Limnophora	6			PR
Lispe	6			PR
Lispoides	6			PR
Phaonia	6			PR
Spilogona	6			PR
Phoridae	**6**			**CG**
Dohrniphora	6			CG
Megaselia	6			CG
Scathophagidae	**6**			**SH**
Acanthocnema	6			SH
Cordilura	6			SH
Hydromyza	6			SH
Orthacheta	6			PR

Spaziphora	6			SC
Syrphidae	**10**			**CG**
Blera	10			CG
Callicera	10			CG
Ceriana	10			CG
Chalcosyrphus	10			CG
Chrysogaster	10			CG
Eristalinus	10			CG
Helophilus	10			CG
Mallota	10			CG
Myolepta	10			CG
Neoascia	10			CG
Sericomyia	10			CG
Spilomyia	10			CG
Tipulidae	**4**	**4**	**4**	**SH**
Brachypremna	4			SH
Leptotarsus	4			SH
Prionocera	4			SH
Tipula	4	5	5	SH
Phalacrocera	4			SH
Triogma	4			SH
Antocha	3	4	4	CG
Arctoconopa	4			SH
Cryptolabis	4	3	3	CG
Dactylolabis	4			SH
Dicranota	3	3	3	PR
Elliptera	4			SH
Gonomyia	4			SH
Helius	4			SH
Hexatoma	2	3	3	PR
Limnophila	3	4	4	PR
Limonia	6	4	4	SH
Molophilus	4	3	3	SH
Ormosia	6	3	3	CG
Paradelphomyia	4			SH
Pedicia	6	3	3	PR
Pilaria	7	4	4	PR
Pseudolimnophila	2	4	4	PR
Rhabdomastix	4			SH
Ulomorpha	4			PR
Erioptera	7	4	4	CG
Lipsothrix	4	4	4	SH
Culicidae	**8**			**FC**

Aedes	8			FC
Anopheles	8			FC
Culex	8			FC
Culiseta	8			FC
Mansonia	8			FC
Orthopodomyia	8			FC
Psorophora	8			PR
Toxorhynchites				PR
Uranotaenia	8			FC
Wyeomyia	8			FC
Simuliidae	**6**			**FC**
Cnephia	4	3	3	FC
Ectemnia	1			FC
Greniera	6			FC
Prosimulium	2	3	3	FC
Simulium	6	5	5	FC
Stegopterna	6			FC
Twinnia	6			FC
Chironomidae	**6**	**5**	**5**	**CG**
Sciomyzidae	**10**			**PR**
Spongillidae	**4**			**FC**
Hydridae	**4**			**PR**
Cavidae	**4**			**PR**
Petasidae	**4**			**PR**
Turbellaria	**9**	**5**	**5**	**PR**
Nemertea	**6**	**4**	**4**	**PR**
Nematoda	**9**			**CG**
Gastropoda				
Valvatidae	**2**	**4**	**4**	**SC**
Viviparidae	**7**	**4**	**4**	**CG**
Ampullaridae	**7**			**SC**
Bithyniidae	**7**			**SC**
Micromelaniidae	**7**			**SC**
Hydrobiidae	**8**	**4**	**4**	**SC**
Pomatiopsidae	**8**			**SC**
Pleuroceridae	**7**	**4**	**4**	**SC**
Lymnaeidae	**7**	**5**	**5**	**SC**
Physidae	**8**	**5**	**5**	**SC**
Planorbidae	**6**	**5**	**5**	**SC**
Ancylidae	**7**	**4**	**4**	**SC**
Margaritiferidae	**5**			**FC**

Unionidae	**4**			**FC**
Sphaeriidae	**8**			**FC**
Corbiculidae	**4**	**5**	**5**	**FC**
Dreissenidae	**5**			**FC**
Hirudinea	**8**	**5**	**5**	**PR**
Oligochaeta	**10**	**5**	**5**	**CG**
Tubificidae	**10**	**5**	**5**	**CG**
Branchiobdellida	**6**	**4**	**4**	**CG**
Polychaeta	**10**			**FC**
Amphipoda	**6**	**4**	**4**	**CG**
Crangonyctidae	**4**			**CG**
Crangonyx	4	4	4	CG
Stygonectes	4			CG
Gammaridae	**4**			**CG**
Gammarus	4	4	4	CG
Haustoriidae	**5**			**CG**
Monoporeia	5			CG

Pontoporeiidae	**5**			**CG**
Hyalella	8	4	4	CG
Decapoda		**4**	**4**	**UK**
Cambaridae	**6**	**4**	**4**	**CG**
Cambarus	6	4	4	CG
Fallicambarus	6			CG
Orconectes	6	4	4	CG
Procambarus	6			SH
Isopoda	**8**	**5**	**5**	**CG**
Asellidae	**8**	**5**	**5**	**CG**
Caecidotea	6	5	5	CG
Lirceus	8	6	6	CG
Ostracoda	**8**			**CG**
Cladocera	**5**			**FC**
Bryozoa	**4**			**FC**
Hydracarina	**7**	**4**	**4**	**PR**
Nematomorpha	**9**			**CG**